Progress of Theoretical Physics, Vol. 20, No. 5, November 1958

New Interpretation of Hyperonic Charge and its Generalization to Leptons*

Takao OKABAYASHI

Department of Physics, University of Tokyo, Tokyo

(Received July 1, 1958)

All weak interactions are assumed to violate the conservation law of parity, contrary to the case of strong interactions. Such correlation between two conservation laws of parity and of hyperonic charge can be achieved by a new interpretation of hyperonic charge. The phenomenological hyperonic charge is expressed by the expectation value of a new operator $\zeta_2\gamma_5$. Nucleon and Ξ-particle are assumed as the eigenstates of this operator, and two τ-spin doublets recombined from Λ-and Σ-particles are assumed as the eigenstates of $\zeta_3\gamma_5$, which anticommutes with $\zeta_2\gamma_5$. If we assign the suitable ζ-dependence to the tensor matrices associated with K- and π-mesons, we can arrive at the correlation mentioned above. In the course of the construction of our scheme, the problems of the so-called particle mixture, and of the particle image under the coexistence of the parity-conserving and nonconserving interactions are used as the stepping stones. The introduction of the new matrix ζ and the ζ-dependence of the tensor matrices associated with bosons allow us to assign the generalized hyperonic charge to leptons, and to exclude the unwanted processes, $\pi\to e+\nu$, $(K\to e+\nu)$ and $P+\mu\to P+e$, owing to the assignment. The connection of our scheme and the interpretation that hyperonic charge can be understood as the parity in isotopic spin space suggests some internal correlation between the Minkowski space and the isotopic spin space. The possibility of such correlation is discussed. In Appendix, the spinor analysis in the six dimensional space, which partly reproduces our phenomenologically obtained results, is described as a tentative model.

§ 1. Introduction

Assigning of isotopic spin and strangeness to baryons and mesons, and classifying of interactions into weak and strong ones[1] have marked an epoch in the study of elementary particles. The strong interactions are characterized clearly by the conservation of strangeness and isotopic spin, and parity is also conserved there. Therefore, it seems to be meaningful to introduce these conserving quantities as new quantum numbers as far as we are concerned with strong interactions, in addition to spin, parity, mass and charge, which are considered usually as the quantum numbers of elementary particles. On the other hand, the weak interactions show that these new quantities have characters quite different from the old quantum numbers of elementary particles. Namely, we assigned strangeness to

---

* The preliminary results were reported at the annual meeting of Japan Physical Society held at Tokyo University, November 1957, and published in the "Soryūshiron-kenkyū" (mimeographed circular in Japanese), **15** (1957), 496; **16** (1957), 130, 261 and **14** (1957), 481.

elementary particles in such a way that the sum of the quantum numbers changes by one unit in the transition, and isotopic spin is not conserved there, too. Further, the violation of the conservation of parity is observed not only in weak Fermi interactions[2], but also in the decay of Λ particle[3], in which neutrino plays no role explicitly. The so-called $\tau-\theta$ puzzle can be understood easily, if the parity conservation is violated in the concerned weak interaction.[4]

The facts that parity and strangeness are conserved in the strong interactions and that neither of them seems to be conserved in all weak interactions*) may be considered to suggest the existence of some correlation between these two conservation laws. Then, we can take up this correlation as a key to understand the singular character of the new quantum numbers. In this article, we want to understand this correlation by introducing a new set of Pauli-type matrices ζ_i's, and reinterpret hyperonic charge, which is the equivalent parameter to strangeness regarding to the selection rule, as the expectation value of the matrix $\zeta_2\gamma_5$. The so-called hyperonic charge is only a kind of phenomenological c-number parameter, but the problem of the particle mixture[5] suggests that it is more promising to understand hyperonic charge as the expectation value of some operator, as will be explained in the next section. On the other hand, hyperonic charge does not always conserve. Therefore, the operator will not be expressed by a simple generating operator associated with any group, as energy-momentum or angular momentum. We take account of this situation by assigning the different operators to baryons with hyperonic charge $\pm$ 1 and the ones with hyperonic charge zero as the constants of motion which are connected to hyperonic charge. It is more convenient to recombine Λ- and Σ-particles into two isotopic spin doublets Y and Z with hyperonic charge zero, as was done in Gell-Mann's[6] article on the grobal symmetry. Nucleon and Ξ-particle with hyperonic charge ± 1 are assumed to be the eigenstates of $\zeta_2\gamma_5$, and Y-and Z-particle are formulated as the eigenstates of the other matrix $\zeta_3\gamma_5$, which anticommute with $\zeta_2\gamma_5$. In this way, we can arrive at a phenomenological scheme, which expresses the correlation of the two conservation laws, if we assign the suitable ζ-dependence to the tensor matrices associated to bosons. The appearance of the matrix γ_5 in the hyperonic charge operator is a reflection of the concerned correlation.

In the next section, we shall introduce the matrix ζ in connection with the modification of the particle image in the case where the parity-conserving and nonconserving interactions exist simultaneously. There the guiding principle is that the clothed free particle should have the same particle image as the bare particle. The fundamental equations of spinor fields are modified so as to regain the desired particle image. The properties of the three equations, which are possible from the correspondence-theoretical point of view, are analyzed. The correspondence of the types of spinor field to physical baryons is determined in such a way that the concerned correlation between the two con-

*) It is not established yet that the conservation of parity is always violated in the weak baryon interactions. For example, this violation is not confirmed for the Σ-decay.[3] But here we assume that parity is not conserved in all weak interactions, at least in the lowest approximation.

servation laws is achieved. In order to reproduce the usual selection rule about hyperonic charge, it is necessary to go through some suitable virtual states for some interactions. But the additional procedure does not disturb the selection rule and the correlation. The situation that it is sufficient for the phenomenological hyperonic charge to get the selection rule between the initial and final states will be emphasized in Appendix I.

The concerned correlation between the two conservation laws may be brought about by some physical entity such as Λ particle in Sakata's composite particle model.[7] But, if we take account of this correlation along the line which will be described in § 2, we can systematically understand not only the phenomena in which only baryons and mesons take parts, but the weak interactions in which leptons also participate. In § 3, we shall assign the types of spinor field to leptons so as to exclude the unwanted processes $\pi \to e+\nu$, $(K \to e+\nu)$ and $P+\mu \to P+e$, and to allow the reactions $\pi \to \mu+\nu$, $K \to \mu+\nu$, μ-capture, $\mu \to e+\nu+\nu$ and β-decay. There we shall assume the conservation of lepton number, and that μ^- is particle. There do not occur any parity-conserving interactions (consequently, strong interactions), when leptons participate. In § 4, few explanations about the idea proposed in this article will be supplied, and the problems which may be comprised in the line of our considerations will be enumerated. Especially, the connection between our scheme and the interpretation that hyperonic charge can be understood as the parity in isotopic spin space will be discussed. It may be emphasized here that our interpretation of hyperonic charge is formulated without reference to isotopic spin space. In Appendix II, a preliminary model of the spinor analysis in the six dimensional space, which leads the modified wave equations, will be analysed, in order to facilitate the understanding of the idea, and to interpret the problems which are not sufficiently explained in the text. Further, it may help us in getting deeper understanding about the parity violation, and in surmising the internal correlation between the isotopic spin space and the Minkowski space.

§ 2. New interpretation of hyperonic charge

Strangeness is really a good dynamical parameter for the phenomena in which the so-called strange particles participate. But hyperonic charge seems to be more suitable, in order to formulate the symmetry character basing on isotopic spin space,[8][9] the mass-multiplicity, the dynamical properties of mesons,[10] and the grobal symmetry.[6] Therefore, in the following, we use hyperonic charge instead of strangeness, and consider the pairs (N, Ξ) with hyperonic charge ± 1 and (Y, Z) with hyperonic charge zero as the physical entities in a more symmetrical world, along the line of Schwinger[10] and Gell-Mann.[6] Here Y and Z are two isotopic spin doublets suitably recombined from Λ- and Σ-particles, and may be defined as $Y=\left(\Sigma^+, \frac{\Lambda^0-\Sigma^0}{\sqrt{2}}\right)$ and $Z=\left(\frac{\Lambda^0+\Sigma^0}{\sqrt{2}}, \Sigma^-\right)$.

We are embarrassed firstly by the problem of particle mixture,[5] when we want to take account of hyperonic charge as the new quantum number of the strange particles.

The θ meson, which decays into two π mesons, is not a θ^0 meson with a definite hyperonic charge 1, but a θ_1^0 meson defined by

$$|\theta_1^0> = \frac{1}{\sqrt{2}}(|\theta^0> + |\tilde{\theta}^0>),$$

if C- or PC-invariance is satisfied in the decay. (Here $\tilde{\theta}^0$ means the antiparticle of a θ^0-meson.) Owing to this situation, the θ^0 meson has two kinds of lifetime, contrary to θ_1^0. If a particle should have its unique lifetime, as supposed by Pais and Gell-Mann, θ_1^0 is considered to be a particle. But, θ_1^0 has not a definite hyperonic charge. Thus there appears a discrepancy between Pais-Gell-Mann's particle image and the one which supposes for a particle to have a definite hyperonic charge.

We can obtain a new particle image by utilizing this dilemma,* if we take into account the fact that θ_1^0 and θ_2^0 have definite parities with respect to the operation of particle-antiparticle conjugation. Namely, it is sufficient to solve this discrepancy to suppose that hyperonic charge is not a c-number but an expectation value of some operator η, and that this operator does not commute with the operator for particle-antiparticle conjugation** A. Then, it will be possible to formulate the theory in such a way that the expectation value of η is observed in the production of θ-meson, and that the parity with respect to A is observed in its decay.

We had to assign the mass multiplets concerning the isotopic spin beforehand when we introduced strangeness. Therefore, it is natural to try to understand this parameter in the frame of isotopic spin space. But this problem is postponed to the end of § 4, and here we use hyperonic charge as a kind of phenomenological parameter describing the usual selection rule, which will be explained in detail in Appendix I. In order to determine the commutation relation between η and A, we can use the fact that we assumed for the hyperonic charge of the antiparticle to have the inverse sign and the same magnitude to the one of the corresponding particle. This situation is formulated as***

$$\{A, \eta\} = 0 . \qquad (2\cdot1)$$

This requirement compels us to modify the wave equations in such a way that the new equations contain some new operator. Because, the operator η is to be understood to contain some operator which does not appear in the conventional wave equation, and the operator A is the one associated with the wave equation.

---

* It might be possible to assume that C- and PC- (consequently T-) invariances fall down in the decay. But there is no evidence until now that T invariance is violated in the nature. Further, the θ_2^0 meson with a long life is really observed.[11]

** The conjugated field to the field φ is expressed by

$$\varphi^c = A\varphi.$$

*** The proof of this equation is as follows. We get

$$\eta\ A|\text{particle}> = \eta|\text{anti-particle}> = -U|\text{anti-particle}>$$

and

$$A\eta|\text{particle}> = UA|\text{particle}> = U|\text{anti-particle}>,$$

for the arbitrary one particle state with hyperonic charge U.

Here we turn our attention, for a little while, to the particle image under the coexistence of the parity-conserving and -nonconserving interactions. If we calculate the self-energy of some particle under such situation, the intrinsic parity of the clothed particle will be different from the one of the corresponding bare particle, generally. Therefore, the equation of motion of the clothed nucleon, with the Fermi-interaction proposed by Lee and Yang,[12] will become

$$a(\gamma_\mu \partial_\mu (1+b\gamma_5)+m')\psi'=0, \qquad (2\cdot 2)$$

where the coefficients a, b and m' are required to be real by the hermitian character of $\bar{\psi}\gamma_\mu \partial_\mu \psi$, $\bar{\psi}\gamma_5\gamma_\mu \partial_\mu \psi$, and $\bar{\psi}\psi$*). The b-term comes from the self-energy graphs, in which odd numbers of parity-nonconserving interactions participate. In fact, the lowest order calculation in the Green function method gives non-zero value to the coefficient b, except for the special types of interactions (see Table I). Now,

Table I.

H_n \ H_c	S	P	V	A	T
S'	B	L	B	L	BL
P'	L	B	L	B	BL
V'	B	L	B	L	
A'	L	B	L	B	
T'	BL	BL			

In this table, such combinations of H_c and H_n that give the zero value to the coefficient b in the equation (2.2) are marked by B for the nucleon self-energy, and by L for the lepton self-energy. H_c and H_n mean the parity-conserving and-nonconserving interactions, respectively, which act at each vertex in the lowest order self-energy graph. The names of the interactions are used after Lee and Yang[12]. The difference between the cases of nucleon and leptons is due to the way that the extra γ_5 matrix enters in the closed loop. It is not essential here that the mass of neutrino vanishes.

$$\gamma_\mu\text{'s and } \frac{1}{\sqrt{1-b^2}}\gamma_\mu(1+b\gamma_5)\text{'s}$$

satisfy the same commutation relation, as will be expected from the fact that we obtain Klein-Gordon's equation from the equation (2.2) by the iteration as in the case of Dirac's equation. Therefore, there exists such transformation V that satisfies

$$\frac{1}{\sqrt{1-b^2}}\gamma_\mu(1+b\gamma_5)=V^{-1}\gamma_\mu V. \qquad (2\cdot 3)$$

We can solve this equation, for $b<1$, and get

$$V=e^{-f\gamma_5}, \qquad (2\cdot 4)$$

where

$$f=\frac{1}{2}\cosh^{-1}\frac{1}{\sqrt{1-b^2}}.$$

That is, the equation (2.2) can be reduced to the usual Dirac's equation

$$(\gamma_\mu \partial_\mu + m_{\text{obs}})\psi_r=0,$$

if we put

* Here, we assumed the T-invariance, for simplicity. This assumption is not essential to the following discussions. If we do not require this invariance, the equation (2·2) must be replaced by

$$a\{\gamma_\mu\partial_\mu(1+b\gamma_5)+m'(1+c\gamma_5)\}\psi'=0. \qquad (2\cdot 2')$$

This equation also can be reduced to Klein-Gordon's equation. But there remains γ_5 term in the mass term, when we perform the transformation (2.3), which is defined uniquely in order to reduce the $\gamma_\mu\partial_\mu(1+b\gamma_5)$ term to the ordinary form $\gamma_\mu\partial_\mu$.

$$\psi_r = \sqrt{a}\ e^{f\gamma_5}\psi'. \tag{2·5}$$

This substitution is a kind of renormalization. The renormalization of the wave operator is the multiplication of the ordinary number, in the parity conserving theory. Here the renormalization contains the multiplication of $e^{f\gamma_5}$ and this factor mixes the two types of field, ψ' and $\gamma_5\psi'$, which have the different intrinsic parity. (In other words, the mixing rate of the large and the small components must be renormalized.) Therefore, it becomes objectionable to the renormalization theory to treat parity as the quantum number of particles, if the parity-nonconserving interactions exist. Further, it is desirable for some quantity to be renormalized to appear in the starting form of the theory, in the renormalization philosophy*.

Taking account of such situations, we propose** to modify the fundamental equation of spinor fields as

$$\{(\alpha+\beta\gamma_5)\gamma_\mu\partial_\mu+m\}\psi=0, \tag{2·6}$$

and suppose that this equation reduces to the ordinary wave equation correspondence-theoretically in the limit, where the parity asymmetry disappears in the nature, (i.e. $\alpha\to 1$ and $\beta\to 0$). It is more convenient for our present purpose to make the equation (2.6) couple with the equivalent one for $\gamma_5\psi$, which has the different intrinsic parity from ψ by the minus sign. Such equation is obtained by multiplying the equation (2.6) by γ_5 from left, as

$$\{(\alpha+\beta\gamma_5)\gamma_\mu\partial_\mu-m\}(\gamma_5\psi)=0. \tag{2·7}$$

Then, we understand these equations as the coupled equation for ψ and $\gamma_5\psi$, or for some two independent linear combinations of them. If we introduce a new set of Pauli's matrices*** ζ_i's $(i=1,2,3)$, this coupled equation can be written in the following three modes.

1) $$\{i(\alpha+\beta\zeta_1)\zeta_2\gamma_5\gamma_\mu\partial_\mu-m\}\Psi_I=0, \tag{2·8–I}$$

$$\Psi_I=\begin{pmatrix}\psi\\ \gamma_5\psi\end{pmatrix}. \tag{2·9–I}$$

2) $$\{i(\alpha+\beta\zeta_2)\zeta_1\gamma_5\gamma_\mu\partial_\mu+m\}\Psi_{II}=0, \tag{2·8–II}$$

---

* In the renormalization theory, it is assumed for the clothed particle to be represented by the same irreducible representation of the Lorentz group as the bare particle.

** It is possible to consider that the existence of the parity-nonconserving interactions has no meaning as far as we treat the clothed but free particles. But the transformation (2.5) produces a weaker parity nonconserving interaction from a parity conserving interaction $\bar{\psi}'O\psi'$, generally. That is,

$$\bar{\psi}'O\psi'\to\bar{\psi}_r\, e^{-f\gamma_5}\, O\, e^{f\gamma_5}\,\psi_r.$$

Of course, $\bar{\psi}'O\psi'$ keeps its form invariant, if O commutes with γ_5.

*** In this article, the representation of ζ_i's is fixed as

$$\zeta_1=\begin{pmatrix}0&1\\1&0\end{pmatrix},\quad \zeta_2=\begin{pmatrix}0&-i\\i&0\end{pmatrix},\quad \zeta_3=\begin{pmatrix}1&0\\0&-1\end{pmatrix}.$$

$$\Psi_{II}=\begin{pmatrix}\psi\\ i\gamma_5\psi\end{pmatrix}. \qquad (2\cdot 9\text{–II})$$

3) $$\{(\beta+\alpha\zeta_3)\zeta_1\gamma_5\gamma_\mu\partial_\mu+m\}\,\Psi_{III}=0, \qquad (2\cdot 8\text{–III})$$

$$\Psi_{III}=\begin{pmatrix}(1+\gamma_5)/2\cdot\psi\\ (1-\gamma_5)/2\cdot\psi\end{pmatrix}. \qquad (2\cdot 9\text{–III})$$

Indeed, if we substitute (2.9) in the equations (2.8) respectively, we obtain the equations (2.6) and (2.7).* Further, the matrices $\zeta_1\gamma_5$, $\zeta_2\gamma_5$ and $\zeta_3\gamma_5$ are the constants of motion in the equations (2.8–I$\sim$III), respectively, and Ψ_I, Ψ_{II} and Ψ_{III} are their eigenstates with the eigenvalue $+1$, as will be seen easily. In the following, we denote them as

$$\zeta_1\gamma_5\,\Psi_I{}^{\pm}=\pm\Psi_I{}^{\pm}=\pm\begin{pmatrix}\psi^{(\pm)}\\ \pm\gamma_5\psi^{(\pm)}\end{pmatrix}, \qquad (2\cdot 10\text{–I})$$

$$\zeta_2\gamma_5\,\Psi_{II}{}^{\pm}=\pm\Psi_{II}{}^{\pm}=\pm\begin{pmatrix}\psi^{(\pm)}\\ \pm i\gamma_5\psi^{(\pm)}\end{pmatrix}, \qquad (2\cdot 10\text{–II})$$

and

$$\zeta_3\gamma_5\,\Psi_{III}{}^{\pm}=\pm\Psi_{III}{}^{\pm}=\pm\begin{pmatrix}(1\pm\gamma_5)/2\cdot\psi^{(\pm)}\\ (1\mp\gamma_5)/2\cdot\psi^{(\pm)}\end{pmatrix}, \qquad (2\cdot 10\text{–III})$$

where $\psi^{(+)}$ satisfies the equations (2.6) and (2.7) and $\psi^{(-)}$ satisfies the similar equations with the opposite signs of β and m. In the following, it is assumed that $\psi^{(-)}$ has the same intrinsic parity as $\psi^{(+)}$**

Now, we may attempt to set up some relation between the operators η and ζ in such a way that the equation (2.1) is satisfied and that we get the correlation between the two conservation laws of parity and of hyperonic charge, which is assumed in the Introduction. The particle-antiparticle conjugation operators for the equations (2.8–I$\sim$III) are given by***

$$A_I=\rho_3\,\sigma_2\,K, \qquad (2\cdot 11\text{–I})$$

$$A_{II}=\zeta_3\,\rho_3\,\sigma_2\,K, \qquad (2\cdot 11\text{–II})$$

---

* Therefore, the three equations in (2.8) are equivalent concerning this point. But each of them plays a different role in the following, when they are considered simultaneously. Further, they are essentially different from each other in the spinor analysis in the six-dimensional space, as will be seen in Appendix II.

** This assumption is necessary to get the results included in this article. At this stage, we can say nothing about the relative parity between $\psi^{(-)}$ and $\psi^{(+)}$, as we have no formal foundation to treat the parameter β yet.

*** With respect to the γ matrix, we used the standard representation, that is,

$$\gamma_4=\rho_3,\ \gamma_k=\rho_2\sigma_k,\ \text{and}\ \gamma_5=-\rho_1.$$

The representation of Weyl-Pauli,

$$\gamma_4=\rho_1,\ \gamma_k=-\rho_2\sigma_k\ \text{and}\ \gamma_5=-\rho_3,$$

gives the same results for the relations (2.12).

and $$A_{III}=i\sigma_3\sigma_2 K, \tag{2·11–III}$$

respectively, where the operation K means to perform complex conjugation. The operators $\zeta_1\gamma_5$, $\zeta_2\gamma_5$ and $\zeta_3\gamma_5$ satisfy the relation (2.1), if we use the corresponding operators A_i, as

$$\{\zeta_1\gamma_5,\ A_I\}=0, \tag{2·12–I}$$

$$\{\zeta_2\gamma_5,\ A_{II}\}=0, \tag{2·12–II}$$

and $$\{\zeta_3\gamma_5,\ A_{III}\}=0. \tag{2·12–III}$$

Therefore, the three equations (2.8) are equivalent, as far as we are concerned with the relation (2.1). If we select out one of the three $\zeta_i\gamma_5$'s, $\zeta_a\gamma_5$, as the operator γ_j, we can describe the nucleon field with hyperonic charge $+1$ by the spinor $\Psi_a{}^+$, and the Ξ field with hyperonic charge -1 by the spinor $\Psi_a{}^-$. As to the Y- and Z-field with hyperonic charge zero, we assume that they are described by the spinors $\Psi_b{}^\pm$, which are the eigenstates of $\zeta_b\gamma_5$ different from $\zeta_a\gamma_5$. Then the expectation values of $\gamma_j=\zeta_a\gamma_5$ in these states become zero*. Thus, we can reexpress hyperonic charge at least in the kinematical feature of baryon fields, along the line proposed in the beginning of this section.

The operator $\zeta_a\gamma_5$, which is to be identified with the hyperonic charge operator γ_j, and $\zeta_b\gamma_5$, which is the constant of motion of Y- and Z-fields, can be selected out of the three $\zeta_i\gamma_5$'s by taking account of the dynamical character of hyperonic charge. In this article we assume that the conservation law of parity is always violated in weak interactions, contrary to strong interactions, as was explained in the previous section. Therefore, we must formulate our scheme in such a way that the following requirements are satisfied: a) parity is conserved in the interactions which conserve hyperonic charge, b) the conservation law of parity is violated in the interactions which change hyperonic charge by one unit, and c) all interactions, which change hyperonic charge by the amount more than two, are forbidden.

The Yukawa type interaction between two spinor fields Ψ and Ψ' and one boson field Ω can be rewritten as

$$(\bar{\Psi}O\Psi')\,\Omega=\mathrm{sp}(\Psi'\,\bar{\Psi}\,G), \tag{2·13}$$

where $G=O\Omega$ and O is some combination of γ-, ζ- and τ-matrices. (In the following, we restrict ourselves to the non-derivative couplings). The operator G is a tensor matrix associated with the boson field Ω. In the conventional field theory, the γ-dependence of the tensor matrix G associated with some boson field is determined by the transformation properties of spinor fields, as will be clearly seen in the so-called fusion

---

* The proof is as follows,

$$\Psi_b{}^{+\dagger}\zeta_a\gamma_5\Psi_b{}^+=1/2\cdot\Psi_b{}^{+\dagger}(\zeta_b\gamma_5\cdot\zeta_a\gamma_5+\zeta_a\gamma_5\cdot\zeta_b\gamma_5)\,\Psi_b{}^+$$

$$=1/2\cdot\Psi_b{}^{+\dagger}(\zeta_b\zeta_a+\zeta_a\zeta_b)\,\Psi_b{}^+=0\,.$$

method.[13] In the present theory, the fundamental equations for spinor fields (2.8) are considered to involve some essentially new elements, and to be formulated as the representation of some group wider than the Lorentz group. But we cannot say anything about the new group at the present stage, and we want to try to determine the ζ-dependence of the tensor matrix G phenomenologically in such a way that our requirements are fufilled. If we can find the suitable ζ-dependence, which is common to all interactions of π-(K-) meson, we may expect that the nature suggests us some possibility to extend the Lorentz group.

In order to analyse the interactions (2.13), we must construct the adjoint field $\overline{\Psi}_i$ to Ψ_i in such a way that they satisfy the corresponding adjoint equations to (2.8–I$\sim$III)

$$\overline{\Psi}_I\{i(\alpha+\beta\zeta_1)\zeta_2\gamma_5\gamma_\mu\partial_\mu-m\}=0, \tag{2·14–I}$$

$$\overline{\Psi}_{II}\{i(\alpha+\beta\zeta_2)\zeta_1\gamma_5\gamma_\mu\partial_\mu+m\}=0, \tag{2.14–II}$$

and

$$\overline{\Psi}_{III}\{(\beta+\alpha\zeta_3)\zeta_1\gamma_5\gamma_\mu\partial_\mu+m\}=0. \tag{2·14–III}$$

These spinors can be obtained easily as

$$\overline{\Psi}_{I\pm}=\Psi_I^{\dagger}\pm\gamma_4\zeta_3=(\overline{\psi}^{(\pm)},\ \pm\overline{\psi}^{(\pm)}\gamma_5), \tag{2·15–I}$$

$$\overline{\Psi}_{II\pm}=\Psi_{II\pm}^{\dagger}\gamma_4\zeta_3=(\overline{\psi}^{(\pm)},\ \mp i\overline{\psi}^{(\pm)}\gamma_5) \tag{2·15–II}$$

and

$$\overline{\Psi}_{III\pm}=\Psi_{III\pm}^{\dagger}\gamma_4\zeta_1=\left(\overline{\psi}^{(\pm)}\frac{1\pm\gamma_5}{2},\ \overline{\psi}^{(\pm)}\frac{1\mp\gamma_5}{2}\right), \tag{2·15–III}$$

Table 2.

	$\overline{\Psi}_I\Phi_I$		$\overline{\Psi}_{II}\Phi_{II}$		$\overline{\Psi}_{III}\Phi_{III}$		$\overline{\Psi}_{II}\Phi_I$		$\overline{\Psi}_{III}\Phi_I$		$\overline{\Psi}_{III\pm}\Phi_{II}{}^{+}$	$\overline{\Psi}_{III\pm}\Phi_{II}{}^{-}$
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)		
1	S	0	S	0	S	0	$(1-i)S$	$(1+i)S$	$\pm PS$	S	$(1\mp i)(S+iPS)$	$(1\pm i)(S-iPS)$
ζ_1	$\pm PS$	0	0	$\mp iPS$	0	S	$(\pm1\mp i)PS$	$(\mp1\mp i)PS$	S	$\mp PS$	$(1\pm i)(S+iPS)$	$(1\mp i)(S-iPS)$
ζ_2	0	$\pm iPS$	$\pm PS$	0	0	$\mp iPS$	$(\pm1\mp i)PS$	$(\pm1\pm i)PS$	$\mp iPS$	iS	$(\pm1+i)(S-iPS)$	$(\mp1+i)(S+iPS)$
ζ_3	0	S	0	S	$\pm PS$	0	$(1+i)S$	$(1-i)S$	S	$\pm PS$	$(1\pm i)(S-iPS)$	$(1\mp i)(S+iPS)$
γ_5	PS	0	PS	0	PS	0	$(1-i)PS$	$(1+i)PS$	$\pm S$	PS	$(\pm1+i)(S-iPS)$	$(\pm1-i)(S+iPS)$
$\gamma_5\zeta_1$	$\pm S$	0	0	$\mp iS$	0	PS	$(\pm1\mp i)S$	$(\mp1\mp i)S$	PS	$\mp S$	$(\mp1+i)(S-iPS)$	$(\mp1-i)(S+iPS)$
$\gamma_5\zeta_2$	0	$\pm iS$	$\pm S$	0	0	$\mp iS$	$(\pm1\mp i)S$	$(\pm1\pm i)S$	$\mp iS$	iPS	$(1\mp i)(S+iPS)$	$(-1\mp i)(S-iPS)$
$\gamma_5\zeta_3$	0	PS	0	PS	$\pm S$	0	$(1+i)PS$	$(1-i)PS$	PS	$\pm S$	$(\pm1-i)(S+iPS)$	$(\pm1+i)(S-iPS)$

In the column (a), the matrix elements for the cases, where the sign of the eigenvalue of Ψ is the same as the one of Φ, are written.

In the column (b), the matrix elements for the cases, where the sign of the eigenvalue of Ψ is different from the one of Φ, are written.

The upper sign added in front of S or PS is to be taken for the case, where the eigenvalue of Ψ is $+1$ and the lower sign is for the another case, and S and PS mean scalar and pseudoscalar, respectively.

respectively. Further, these spinors have the properties,

$$\overline{\Psi}_I \pm \zeta_1 \gamma_5 = \pm \overline{\Psi}_I \pm, \tag{2.16-I}$$

$$\overline{\Psi}_{II} \pm \zeta_2 \gamma_5 = \pm \overline{\Psi}_{II} \pm, \tag{2·16-II}$$

and

$$\overline{\Psi}_{III} \pm \zeta_3 \gamma_5 = \pm \overline{\Psi}_{III} \pm, \tag{2·16-III}$$

as will be easily seen by their definition (2.15). The matrix elements between six paris of spinors $\overline{\Psi}_i$ and $\Psi_j (i, j = I, II, III)$ are tabulated in table II, for ζ_k's and $\gamma_5 \zeta_k$'s.

For simplicity, we fix the types of spinor fields for baryons beforehand, and determine the ζ-dependences of the tensor matrices G_π, G_K, and $G_{\tilde{K}}$ for π-, K- and anti-K-mesons, respectively. It will be easily found that there is no other correspondence which satisfies the requirement about the dynamical characters of hyperonic charge. The sole possible correspondence is given by*

$$\Psi_N = \Psi_{II}{}^+, \quad \Psi_\Xi = \Psi_{II}{}^-,$$

and

$$\Psi_Y = \Psi_{III}{}^+, \quad \Psi_Z = \Psi_{III}{}^-. \tag{2·17}$$

In order to find the desirable ζ-dependence of G_π, we first consider the case in which the concerned two spinor fields are Ψ_N and Ψ_Ξ. We must exclude the interactions $(\overline{\Psi}_N O \Psi_\Xi)\pi$ and $(\overline{\Psi}_\Xi O \Psi_N)\pi$, which change hyperonic charge by ± 2. This requirement is satisfied if we choose the combination of 1, ζ_2, γ_5, and $\zeta_2\gamma_5$, which commute with $\zeta_2\gamma_5$, as the matrix O, that is, the ζ-dependence of G_π. Among these matrices, 1 and $\zeta_2\gamma_5$ gives only scalar the biproducts $(\overline{\Psi}_N O \Psi_N)$ and $(\overline{\Psi}_\Xi O \Psi_\Xi)$ as will be seen by Table II. Therefore, the ζ-dependence of G_π must be

$$G_\pi = x\zeta_2 + y\gamma_5 . \tag{2·18}$$

Secondly, we study the biproducts between Ψ_Y and Ψ_Z, all of which conserve hyperonic charge. The matrix ζ_2 anti-commutes with $\zeta_3\gamma_5$, and gives the contributions only to $(\overline{\Psi}_{III} \pm O \Psi_{III} \mp)$, namely $(\overline{\Psi}_Y O \Psi_Z)$ and $(\overline{\Psi}_Z O \Psi_Y)$. The matrix γ_5 commutes with $\zeta_3\gamma_5$, and gives the non-vanishing contributions to $(\overline{\Psi}_{III} \pm O \Psi_{III} \pm)$, namely $(\overline{\Psi}_Y O \Psi_Y)$ and $(\overline{\Psi}_Z O \Psi_Z)$. Both of them are pseudoscalar. At last, parity must not be conserved in the interactions $(\overline{\Psi}_{III} O \Psi_{II})\pi$, which change hyperonic charge by ± 1. The contributions from ζ_2 and γ_5 do not conserve parity generally, as will be seen from Table II. If $x = -y$, π-meson cannot couple with nucleon, and the couplings between Ξ-particle and π-meson are excluded for the case $x = y$. The coefficients x and y are not determined, as far as we are concerned with the interactions between baryons and mesons, except for $x \neq \pm y$, contrary to G_K and $G_{\tilde{K}}$. But in the next section we shall see that it is preferable to be $x = \pm y$ for lepton problems. Therefore, we suppose to put $x = y$ in order to keep the direct $NN\pi$ interaction, and make the interactions of π-meson with Ξ particle go through the

* For instance, the equation (2·8-II) has four independent so'utions, apart from the spin assignment. They correspond to N, $\tilde{N}$, Ξ and $\tilde{\Xi}$ states, accoding to the signs of energy and $\zeta_2\gamma_5$.

virtual states. For instance,

$$\Xi \Rightarrow \tilde{K}+\Lambda \Rightarrow \tilde{K}+(\Lambda+\pi) \Rightarrow \Xi+\pi\,,$$
$$\text{and} \qquad\qquad (2\cdot 19)$$
$$\Xi \Rightarrow \tilde{K}+\Lambda \Rightarrow (\tilde{N}+\Lambda)+\Lambda \rightarrow \pi+\Lambda\,,$$

where $\Rightarrow$ and $\rightarrow$ mean the transitions by strong- and weak-interactions, respectively. This situation will give rise to the splitting between Ξ- and N-levels, and it will also make the Y-level shift from the Z-level indirectly. The splitting between (N, Ξ) and (Y, Z) will be induced by the asymmetry of their weak interactions with the K-meson, as will be given in the following.

The ζ-dependence of G_K and $G_{\tilde{K}}$ can be determined in the following way. A nucleon (Ξ particle) can emit a K- (anti-K-) meson in the transitions to Y or Z, but cannot emit anti-K- (K-) meson. Such situations can be taken into account by making G_K and $G_{\tilde{K}}$ include the projection operator $(1+\zeta_2\gamma_5)/2$ and $(1-\zeta_2\gamma_5)/2$, respectively. Further, these interactions must conserve parity. From these requirements, G_K and $G_{\tilde{K}}$ can be determined as

$$G_K^{\dagger} \sim (\zeta_1-\zeta_3)\frac{1+\zeta_2\gamma_5}{2} \qquad (2\cdot 20)$$

and

$$G_{\tilde{K}}^{\dagger} \sim (\zeta_1-\zeta_3)\frac{1-\zeta_2\gamma_5}{2}, \qquad (2\cdot 21)$$

as will be checked by Table II. Here, for definiteness, we supposed that K- and anti-K-mesons are the pseudoscalar quantities, and that the four-spinors ψ's associated with N, Ξ, Λ, and Σ have the same intrinsic parity. Next, K- and anti-K-mesons can couple with the biproducts between Ψ_Y and Ψ_Z, and the interactions change hyperonic charge by ± 1. Therefore, these interactions must not conserve partiy. In fact, G_K and $G_{\tilde{K}}$ contain four matrices ζ_1, ζ_3, $\zeta_1\gamma_5$ and $\zeta_3\gamma_5$, and the coexistence of these matrices mixes the scalar term and the pseudoscalar term, as will be seen from Table II. At last, we check the interactions with the pairs of Ψ_N and Ψ_Ξ. The interactions $(\overline{\Psi}_N O \Psi_\Xi) K^{\dagger}$ and $(\overline{\Psi}_\Xi O \Psi_N) \tilde{K}^{\dagger}$, which change hyperonic charge by ± 3, are excluded owing to the projection operators. Our definitions (2.20) and (2.21) mix scalar and pseudoscalar for interactions $(\overline{\Psi}_\Xi O \Psi_N) K^{\dagger}$ and $(\overline{\Psi}_N O \Psi_\Xi) \tilde{K}^{\dagger}$. The remaining four direct interactions $(\overline{\Psi}_{II\pm} O \Psi_{II\pm}) K^{\dagger}$ and $(\overline{\Psi}_{II\pm} O \Psi_{II\pm}) \tilde{K}^{\dagger}$, which change hyperonic charge by ± 1, are excluded in our scheme. In order to reproduce these transitions, we must go through the suitable virtual processes. For instance,*

---

* In the last step of (2.22), we used the interaction $\tilde{N}N\pi$. Here we may use the $KK\pi$ interaction similar to the one proposed by Schwinger.[10] In order to construct the parity conserving $KK\pi$ interaction, which conserves hyperonic charge, we must consider the scalar K'-meson. The tensor matrices associated with K'- and $\tilde{K}'$-mesons must be taken as

$$G_{K'}^{\dagger} \sim (\zeta_1+\zeta_3)(1+\zeta_2\gamma_5)/2 \qquad (2\cdot 20')$$
$$G_{\tilde{K}'}^{\dagger} \sim (\zeta_1+\zeta_3)(1-\zeta_2\gamma_5)/2 \qquad (2\cdot 21')$$

$$N \Rightarrow \Lambda + K \rightarrow (N+\pi) + K \Rightarrow N + K . \tag{2·22}$$

Thus we have succeeded in regaining the dynamical character of hyperonic charge by the assignments of (2.17), (2.20) and (2.21). Further, the tensor matrices G_π, G_K and $G_{\tilde{K}}$ have the following properties,

$$\zeta_2 \gamma_5 G_\pi = G_\pi \zeta_2 \gamma_5, \tag{2·23}$$

$$\zeta_2 \gamma_5 G_K = -G_K \zeta_2 \gamma_5 = G_K, \tag{2·24}$$

and

$$\zeta_2 \gamma_5 G_{\tilde{K}} = -G_{\tilde{K}} \zeta_2 \gamma_5 = -G_{\tilde{K}}. \tag{2·25}$$

We may define the hyperonic charge of boson as a, when

$$\zeta_2 \gamma_5 G = -G \zeta_2 \gamma_5 = aG \tag{2·26}$$

and as zero, when

$$\zeta_2 \gamma_5 G = G \zeta_2 \gamma_5 , \tag{2·27}$$

after the methed to define chirality by Watanabe.[13] Then, the assignment coincides with the usual one of hyperonic charge and the problem of particle mixture, which was picked up in the beginning of this section, is reflected in the difference of the sign in the projection operators of G_K and $G_{\tilde{K}}$. We may call attention to the point that the structure of G_π, G_K and $G_{\tilde{K}}$ makes us expect to get the deeper understanding on the new equations (2·8). It will be explained in more detail in Appendix II, by means of a tentative model for the extended six-dimensional space.

§ 3. Lepton problems

We arrived at the equation (2.2), starting from the coexistence of the parity-conserving and -nonconserving interactions, and interpreted the matrix ζ in connection with hyperonic charge. The coefficient b is different from zero generally, except for the cases of the special types of interactions. The same situation occurs for the self-energy of leptons (see Table I), and we must consider that the wave equations of leptons also have the same forms as the equations (2·8). Therefore, we must consider that leptons also have some quantity which has the character of hyperonic charge, notwithstanding we have not still succeeded in the assignment of hyperonic charge or isotopic spin for leptons, contrary to the case of baryons and mesons.

Before entering into the assignment of the generalized hyperonic charge for leptons, a remark on neutrino may be added here. In the case of neutrino, ζ's, in addition to $\zeta_i \gamma_5$'s, are the constants of motion in the equations (2.8), as there is no mass term. Therefore γ_5 becomes the constant of motion. For instance, if we pick up the eigenstates with the eigenvalue $+1$ of ζ_4's from (2.9), the properties

$$\zeta_1 \psi_{I}^{(\pm)} = +\psi_{I}^{(\pm)}, \tag{3·1 I}$$

$$\zeta_2 \psi_{II}^{(\pm)} = +\psi_{II}^{(\pm)}, \tag{3·1 II}$$

and $$\zeta_3 \Psi_{III}{}^{\pm(+)} = +\Psi_{III}{}^{\pm(+)} \qquad (3\cdot 1\text{–III})$$

give $$\gamma_5 \psi = \pm \psi\,, \qquad (3\cdot 2)$$

for the 1st and 2nd types of spinor, and for the 3rd type of spinor we have the structure

$$\Psi_{III}{}^{\pm(+)} = \begin{pmatrix} (1\pm\gamma_5)/2\cdot\psi \\ 0 \end{pmatrix}. \qquad (3\cdot 3)$$

Therefore neutrino has the definite spirality as in the two component theory.[14] Owing to this property, all interactions, in which neutrino plays a role, give the maximum violation of the conservation of parity, regardless of the type of spinor field assigned to neutrino.

In order to assign the generalized hyperonic charge to leptons, we assume the conservation law of lepton number. Further, for μ-meson, we assume μ^- to be particle,* taking into account the analysis on Michel's ρ-parameter.[14] The types of spinor fields associated with leptons can be determined in such a way that the assignment excludes the direct interactions for the so-called unwanted processes.[15]

First we take up the fact that π-meson does not decay into $(e,\ \nu)$**. The direct interaction for the decay can be excluded, if we assign Ψ_{II}- to the spinor of electron Ψ_e. This is due to the projection operator $\gamma_5(1+\zeta_2\gamma_5)/2$ in G_π and to the form of the concerned interaction described in the footnote. If we require to forbid the direct interaction*** for the reaction $P+\mu\to e+P$, we can determine the type of Ψ_μ as $\Psi_{II}{}^+$. The reason is as follows. The concerned Fermi interaction is $(\overline{\Psi}_I O \Psi_P)\,(\overline{\Psi}_e O \Psi_\mu)$, where the matrices, which are inserted between the baryon pair and the lepton pair, are expected to be the same in the future theory (see Appendix II). The ζ-dependences of O which give the non-vanishing contribution to $(\overline{\Psi}_P O \Psi_P)$ must be the ones which commute with $\zeta_2\gamma_5$, as Ψ_P is the eigenstate of $\zeta_2\gamma_5$. If we put $\Psi_\mu = \Psi_{II}{}^+$, such matrices make the matrix element $(\overline{\Psi}_e O \Psi_\mu)$ vanish owing to the previously given assignment $\Psi_e = \Psi_{II}{}^-$. By a similar consideration, the existence of the β-decay excludes the possibility $\Psi_\nu = \Psi_{II}{}^+$, and the μ-capture process excludes $\Psi_\nu = \Psi_{II}{}^-$. Therefore Ψ_ν must be Ψ_I or Ψ_{III}. As to Ψ_ν, we

---

* Then, the concerned direct interactions have the following forms,

$(\overline{\Psi}_e G_K{}^\dagger \Psi_\nu)$ for $K^+\to e^+ + \nu$

$(\overline{\Psi}_\nu G_{\tilde{K}}{}^\dagger \Psi_e)$ for $K^-\to e^- + \tilde{\nu}$

$(\overline{\Psi}_\mu G_K{}^\dagger \Psi_\nu)$ for $K^+\to \mu^+ + \nu$

and $(\overline{\Psi}_\nu G_{\tilde{K}}{}^\dagger \Psi_\mu)$ for $K^-\to \mu^- + \tilde{\nu}$.

If we assume μ^+ to be particle, we have

$(\overline{\Psi}_\nu G_K{}^\dagger \Psi_\mu)$ for $K^+\to \mu^+ + \tilde{\nu}$

$(\overline{\Psi}_\mu G_{\tilde{K}} \Psi_\nu)$ for $K^-\to \mu^- + \nu$.

The similar interactions for π-meson decays are obtained by replacing G_K and $G_{\tilde{K}}$ by G_{π^+} and G_{π^-}, respectively.

** We did not discuss the decay through baryon loops, and it will be necessary to study in more detail. But there remains no question, if we use the V-A coupling.[16]

*** The interactions for μ-capture and for β-decay give rise to the concerned reaction in the second order, but we restrict ourselves to the direct interaction in this article.

put $\Psi_\nu = \Psi_{III}$, on account of the parallelism to the case of baryons and of the economy.* Further, our assignment allows the direct interactions for the decays $\pi \to \mu + \nu$, $\mu \to e - \nu - \nu$ and $K \to e + \nu$, but makes the direct interaction for $K \to \mu + \nu$ vanish. The last decay is obtained through $(\bar{\Psi}_\Lambda O \Psi_I)(\bar{\Psi}_\mu O \Psi_\nu)$, and the strong interaction $(\bar{\Psi}_\Lambda O \Psi_N) G_K$.

If we want to call the expectation value of $\zeta_2 \tilde{\gamma}_5$ the generalized hyperonic charge for leptons, as in the case of baryons, electron, μ^--meson and neutrino have the charge -1, $+1$ and zero, respectively. But, there is no simple selection rule between the generalized hyperonic charge as the one for baryons and mesons. Further, we must distinguish the generalized hyperonic charge of leptons from the hyperonic charge of baryons and mesons. For instance, the decay process $\pi^- \to \mu^- + \tilde{\nu}$ has the different change of the parameter from the one in the process $K^- \to \mu^- + \tilde{\nu}$, if we do not distinguish the two kinds of parameters.

There may be considered another possibility, $\Psi_\nu = \Psi_{II}-$, in order to exclude the direct (e, ν)-decay interaction of π-meson. But such assignment is not relevant. Ψ_ν must not be $\Psi_{II}+$ for the existence of the interaction, which describes the μ-capture process. If we put $\Psi_\mu = \Psi_{II}-$, the absence of the direct interaction for $P + \mu \to P + e$ requires $\Psi_e = \Psi_{II}-$, and then the existence of the β-decay interaction excludes the possibility $\Psi_\nu = \Psi_{II}-$. There remain two possibilities, $\Psi_\mu = \Psi_e = \Psi_I+$ or Ψ_I- and $O = \zeta_2$ or $\zeta_2 \tilde{\gamma}_5$, and $\Psi_\mu = \Psi_I \pm$, $\Psi_e = \Psi_I \mp$ and $O = 1$ or γ_5, in order to exclude the reaction $P + \mu^- \to P + e^-$, as will be easily seen from Table II. But these possibilities depend on the types of interactions. Further, the entrance of Ψ_I is not preferable, as mentioned before.

It may be remarked that we can exclude also the direct interaction for the (e, ν)-decay of K-meson if we put $x = -y$ in the equation (18). If we assign as $\Psi_e = \Psi_{II}+$, $\Psi_\mu = \Psi_{II}-$ and $\Psi_\nu = \Psi_{III}$, the direct interactions for the decays $K \to e + \nu$, $\pi \to e + \nu$, and for the reaction $P + \mu \to P + e$ vanish, and the one for the decays $K \to \mu + \nu$, $\pi \to \mu + \nu$, $\mu \to e + \nu + \nu$, $N \to P + e^- + \tilde{\nu}$, and for the μ-capture are allowed. However, we must make the couplings of π-meson with nucleon go through the suitable virtual states in this case.

If we allow μ^+ meson to be particle, the reaction $P + \mu^- \to P + e^-$ is forbidden owing to the conservation law of lepton number. Therefore, there remain many possibilities. There are two kinds of assignments to exclude the direct interaction for $\pi \to e + \nu$, as in the case where μ^- is particle; (a) $\Psi_e = \Psi_{II}-$, and (b) $\Psi_\nu = \Psi_{II}-$. First, we study the assignment (a). The existence of β-decay excludes $\Psi_\nu = \Psi_{II}-$. The μ-meson capture, $P + \mu^- \to N + \tilde{\nu}$, requires that $\Psi_\nu \neq \Psi_{II}-$ if $\Psi_\mu = \Psi_{II}+$. Therefore, we must have $\Psi_\nu = \Psi_I$ or Ψ_{III} if $\Psi_\mu = \Psi_{II}+$, and $\Psi_\nu = \Psi_{II}-$ or Ψ_I or Ψ_{III}, if $\Psi_\mu = \Psi_{II}-$. The possibilities $\Psi_\mu = \Psi_{II}-$ and $\Psi_\nu = \Psi_{II}-$ make the direct interaction for μ-decay of π-meson vanish, and $\Psi_\mu = \Psi_{II}$ excludes the direct interaction for the (μ, ν)-decay of K-meson, and then these decays are obtained through $(\bar{\Psi}_\nu O \Psi_I)(\Psi_\mu O \Psi_\nu)$ and $(\Psi_\nu O \Psi_I)(\Psi_\mu O \Psi_\nu)$, respectively. The existence of $\mu^- \to e^- + \nu + \tilde{\nu}$ does not give any new restriction. The direct interaction

* As will be seen in Appendix II, it seems to be difficult to construct the spinor analysis in such a way that all equations of (2.8) can be introduced simultaneously.

for (e, ν)-decay of K-meson is not excluded, except for the case $\Psi_\nu = \Psi_{II}^-$. Further, $\Psi_\mu = \Psi_I$ or Ψ_{III} is not excluded. Here, we must supplement a remark that such combinations as $\Psi_\mu = \Psi_\nu = \Psi_I^\pm$ and $O = \zeta_2$ or $\zeta_2\gamma_5$, and $\Psi_\nu = \Psi_I^\pm$, $\Psi_\mu = \Psi_I^\mp$ and $O = I$ or γ_5, make the interaction for the μ-meson capture vanish, as will be seen from Table II. Secondly, we consider the assignment (b). In this case, the direct interaction for the (e, ν)-decay of K-meson is also excluded. The existence of the interaction $P + \mu^- \to N + \tilde{\nu}$ excludes $\Psi_\mu = \Psi_{II}^-$. The existence of β-decay requires $\Psi_e \neq \Psi_{II}^+$. The direct interaction for $\mu^+ \to e^+ + \nu + \nu$ does not give any new restriction. The direct interaction for the (μ, ν)-decay of π-meson is always forbidden. Further, the direct $K \to \mu + \nu$ interaction vanishes, if $\Psi_\mu = \Psi_{II}^-$. But these decays can occur through the baryon loops. Therefore, there remain many possibilities, namely, $\Psi_\mu = \Psi_{II}^-$ or Ψ_I or Ψ_{III}, and $\Psi_e = \Psi_{II}^-$ or Ψ_I or Ψ_{III}.

Summarizing the above analysis, we can assign the types of spinor fields to leptons so as to exclude the direct interactions describing $\pi \to e + \nu$, $P + \mu \to P + e$ (and $K \to e + \nu$) and to allow the direct or indirect reactions, $K \to \mu + \nu$, $\pi \to \mu + \nu$, $\mu \to e + \nu + \nu$, μ-capture and β-decay. The most preferable assignment is to assume μ^- to be particle and $\Psi_e = \Psi_{II}^-$, $\Psi_\mu = \Psi_{II}^+$ and $\Psi_\nu = \Psi_{III}$. If we allow μ^+ to be particle, there are many possibilities; 1) $\Psi_e = \Psi_{II}^-$, $\Psi_\mu = \Psi_\nu = \Psi_{III}$, 2) $\Psi_e = \Psi_{II}^-$, $\Psi_\mu = \Psi_{II}^+$ or Ψ_{II}^- and $\Psi_\nu = \Psi_{III}$, 3) $\Psi_e = \Psi_\mu = \Psi_\nu = \Psi_{II}^-$, 4) $\Psi_e = \Psi_{III}$ and $\Psi_\mu = \Psi_\nu = \Psi_{II}^-$, 5) $\Psi_e = \Psi_\mu = \Psi_{III}$ and $\Psi_\nu = \Psi_{II}^-$, and 6) $\Psi_e = \Psi_\nu = \Psi_{II}^-$ and $\Psi_\mu = \Psi_{III}$. The Ψ_I type spinor field is omitted.

Here we may add a remark on the types of the Fermi interactions. There arises the possibility to classify the interactions into two groups, the (STP)-group and the (V A)-group, owing to the ζ matrices which appear between two spinor fields. For instance, in the case where the concerned two spinor fields are of Ψ_{II} type, there remain only (S, T, P) interactions in $(\bar{\Psi}\, 1 \cdot O_i\, \Psi)$ and $(\bar{\Psi}\, \zeta_2 \cdot O_i\, \Psi)$, and only (A, V) interactions in $(\bar{\Psi}\, \zeta_3\, O_i\, \Psi)$ and $(\bar{\Psi}\, \zeta_1\, O_i\, \Psi)$, as will be easily seen from the commutation properties between $\zeta_j O_i$'s and $\zeta_2\gamma_5$. Here, O_i's are the usual five γ matrices. Further, it may be noteworthy that we cannot construct the parity conserving (consequently strong) interactions in which leptons participate. This is due to the conservation law of lepton number and our assignment.

§ 4. Discussions

In § 2, we have succeeded to correlate the two conservation laws of hyperonic charge and of parity. We had to go a roundabout way to get the same selection rule as the one of the phenomenological hyperonic charge. Namely, we had to go through the virtual loop to arrive at the reactions like (2.19) when we put $x = +y$, or through a self-energy graph for (2.22). But such additional procedures are consistent with the selection rule and the correlation, and it is sufficient to get the usual selection rule for the comparison between the initial and final states, as will be seen from Appendix I. Therefore, our new interpretation of hyperonic charge makes it possible to unify the two characteristic features in strong and weak interactions, namely hyperonic charge and parity.

Further, the interpretation has the merit that it describes many problems inclusively which are up to now treated separately. Firstly, we can get a particle image which gives us a consistent interpretation about the problem of particle mixture. Secondly, the new equations (2.8) include the modification of the particle image under the influence of the parity nonconserving interactions. They will help us in obtaining the deeper understanding about the problem of the parity-violation. At the present stage, we know nothing about the reason why the coupling constants of the parity-nonconserving interactions are weak, and here we assumed *a priori* that the interactions which conserve parity are strong. A clue to solve this problem might be obtained, when we can formulate the theory concretely under some more fundamental requirement in such a way that our ideas explained in the previous sections are included. Another significant merits obtained by the introduction of ζ-matrix are that we can select out the direct interactions for the (e, ν) decay of π (and K-) mesons and the reaction $P+\mu \to P+e$, and that we can assign the generalized hyperonic charge to leptons, as were shown in § 3. Further, the Fermi interactions can be classified into the (STP)-group and the (VA)-group, and the strong interactions in which leptons participate are excluded.

In the remaining of this section, we want to supplement a few explanations about the idea proposed in this article, and to enumerate the problems which may be comprised in the line of our considerations.

A) We assigned the different operators to (Ξ, N) and (Y, Z) as the constants of motion, which are connected to hyperonic charge, that is, $\zeta_{\beta'\beta}$ and $\zeta_{\beta'\beta}$ respectively. Such situation seems to be very complicated from the point of view that we want to formulate hyperonic charge as a quantum number of particles in some sense. But the fact that hyperonic charge is not always conserved seems to show the difficulty to assign a common hyperonic charge operator to all baryons as for the absolutely conserved quantity. Hyperonic charge is only a kind of phenomenological parameter, and it seems to be too simple to express it by a simple operator. The facts that we can obtain the satisfactory assignment of the generalized hyperonic charge for leptons also by $\zeta_{\beta'\beta}$ and $\zeta_{\beta'\beta}$, and that the tensor matrix associated with bosons are characterized by $\zeta_{\beta'\beta}$ make us hope to get a deeper understanding of our scheme.

The situation that the mass multiplets come out due to the shift from the more symmetrical world may be sketched inversely as follows. The (Ξ, N) and (Y, Z) pairs are unified in the equations (2.8–II) and (2.8 III), respectively. Further, both of them correspond to the same coupled equation (2.6) and (2.7).

B) At first sight, it seems to be queer for γ_5 to appear in the operator γ_μ. But there is no reason to exclude the appearance of γ_5 under the interpretation of hyperonic charge as the expectation value. On the contrary, it will be seen that the γ_5 matrix is necessary in order to formulate the correlation between the conservation laws of parity and of hyperonic charge. In the text we used the equations (2.6) and (2.7) as the correspondence-theoretical base, but we may consider such simpler form as

$$(\gamma_\mu \partial_\mu + m)\psi' = 0$$

$$(\gamma_\mu\partial_\mu - m)(\gamma_5\psi) = 0, \tag{4·1}$$

instead of them. Then it means that we start from the wave equations for the elementary waves ψ and $\gamma_5\psi$ composing the wave packet (2.5).* The equation (4.1) can be written as

$$(\gamma_\mu\partial_\mu + m\zeta_3)\,\Psi = 0. \tag{4·2}$$

Here, let us consider more general forms** as

$$(\gamma_\mu\partial_\mu + im\,\zeta_i\gamma_5)\,\Psi = 0, \tag{4·3}$$

Then, we may follow the idea proposed in § 2 by taking a suitable ζ_a as hyperonic charge operator. If we can attain our purpose by describing N and Ξ as the eigenstates of the ζ_a and Y and Z as the ones of some other ζ_b, $(b\neq a)$, we shall get a simpler scheme. The solutions of the three equations (4.3) which are the eigenstates of ζ_i are given, after simple calculations, as

$$\Psi_1 = \begin{pmatrix} \psi \\ \psi \end{pmatrix} \text{and} \begin{pmatrix} \psi' \\ -\psi' \end{pmatrix} \quad \text{for } \zeta_i = \zeta_1\,; \tag{4·4}$$

$$\Psi_2 = \begin{pmatrix} \psi \\ i\psi \end{pmatrix} \text{and} \begin{pmatrix} \psi' \\ -i\psi' \end{pmatrix} \quad \text{for } \zeta_i = \zeta_2, \tag{4·5}$$

and

$$\Psi_3 = \begin{pmatrix} \psi \\ 0 \end{pmatrix} \text{and} \begin{pmatrix} 0 \\ \psi' \end{pmatrix} \quad \text{for } \zeta_i = \zeta_3, \tag{4·6}$$

respectively. Here, ψ and ψ' are the solutions of the equations

$$(i\gamma_5\gamma_\mu\partial_\mu - m)\,\psi = 0$$

and $\tag{4·7}$

$$(i\gamma_5\gamma_\mu\partial_\mu + m)\,\psi' = 0.$$

By any two of Ψ_i's in equations (4.4∼6), we cannot obtain the desired scheme, as will be easily checked by repeating the discussions exhibited in § 2. Further, there are some mathematical defects in such equations. Both ζ_i and $\zeta_j\gamma_5\,(j\neq i)$ are constants of motion in the equation, but these two operators anticommute with each other. Secondly, there are two kinds of particle-antiparticle conjugation operator.

C) The appearance of ζ-matrix in the wave equation will induce many modifications in the field theory. We proposed in 1953[17] to modify the wave equation of nucleon

---

* The renormalization of the wave operator can be interpreted as constructing a kind of wave packet which corresponds to the observed wave. There, it is assumed implicitly that the elementary waves of the packet satisfy the same wave equation (or have the same propagation character). If we start from the coupled equation (4·1), the requirement is satisfied.

** We added γ_5 to the second term in this equation, in order to make the five matrices γ_μ and $\zeta_i\gamma_5$ anticommute with each other. The following discussions are not changed, if this γ_5 factor is omitted

from the consideration about the anomalous magnetic moment of nucleon. We could not formulate the idea concretely at the time, owing to the insufficiency of the knowledge about isotopic spin space. But now the new matrix ζ has been introduced in the fundamental equation. We may expect to arrive at the desired modification, when we introduce the electromagnetic interaction into our equation.

Further, the situation about the well-known divergence problem might be improved owing to the new degree of freedom ζ, which is connected to the Minkowski space, as was emphasized in the previous paper. In Dirac's hole theory, the situation was improved by the passage from the non-relativistic theory. There the matrix ρ plays an important role.

D) The conservation law of hyperonic charge is formulated in isotopic spin space by many authors.[8)9)] Then, hyperonic charge U can be connected to the parity ω in the space* as

$$\omega = e^{i(\pi/2)U}. \tag{4·8}$$

Our selection rule is shifted from the one obtained by them, as we must perform the additional procedures mentioned in the beginning of this section. But both selection rules are almost the same. Therefore, we may try to connect the two interpretations. Of course, we may take the other standpoint that hyperonic charge has no connection with isotopic spin space. Because, we have obtained the equivalent selection rule by our new interpretation, without reference to the space. But, there remains an important problem that the electric charge is given by

$$Q = I_3 + \frac{U}{2}, \tag{4·9}$$

where I_3 is the third component of isotopic spin of particle. This problem might be solved in connection with the unification of the two kinds of charge concepts. Namely, at the present stage, we characterize the charge for some particles by the complex property of the wave function on the one hand, and for some other particles by introducing isotopic spin multiplets on the other hand. Further, if we enlarge isotopic spin space to the four-dimensional one in such a mode as will be given in the following (eq. (4·12)), ζ matrix is embedded in the enlarged space. Therefore, if we impose some invariance character in the space on the whole theory, the behaviour of ζ matrix will automatically give some restriction to the change of the three-dimensional isotopic spin in weak interactions**. Moreover, it is interesting from the purely academic standpoint to explore some connection between the Minkowski space and isotopic spin space. Namely, if there is some internal correlation between the two spaces, we shall be able to get the systematic under-

* Therefore, if we succeed to connect our scheme to the one of d'Espagnat and Prentki, it will suggest the possibility to formulate the whole theory so as to be invariant when we perform the inversions in the Minkowski space and in isotopic spin space simultaneously.

** Even in this case, it will be difficult to give the ratio between the interactions with $|\Delta I| = 1/2$ and $|\Delta I| = 3/2$.[18)]

standing of the quantum numbers of elementary particles inclusive of charge, on the pattern of the connection between the particle image and the space-time structure in the conventional field theory.

The equation (4.8) can be rewritten as the operator equation which operates on the spinors in istopic spin space, as

$$P_\tau = e^{i(\pi/2)\eta'}, \tag{4·10}$$

where P_τ means the inversion operator in the space, and η' is the operator whose expectation value gives hyperonic charge. If we identify the operator η' with $\eta = \zeta_2\gamma_5$, then we get

$$P_\tau = e^{i(\pi/2)\zeta_2\gamma_5} = i\zeta_2\gamma_5. \tag{4·11}$$

This relation suggests to extend isotopic spin space to the four-dimensional one*, in which the reflection matrices are represented as

$$\begin{aligned} T_i &= \zeta_2\tau_i \quad (i=1,2,3), \\ T_4 &= \zeta_j \quad (j\neq 2). \end{aligned} \tag{4·12}$$

Because, the inversion of the three-dimensional τ space embedded in the enlarged space is expressed as

$$P_\tau = T_1T_2T_3 = i\zeta_2,$$

and this coincides with the equation (4.11) apart from the γ_5 matrix, which is equivalent to unit matrix in isotopic spin space.

Such an extension of isotopic spin space requires** the recombination from Λ and $\sum$ to Y and Z, and here the departure from the scheme of d'Espagnat and Prentki occurs. The part of the K-meson interactions in their Lagrangian[8]

$$\bar{N}K\Lambda + \bar{N}\vec{\tau}\overrightarrow{\sum}K + \bar{\Xi}\tau_2K^*\Lambda + \bar{\Xi}\vec{\tau}\overrightarrow{\sum}\tau_2K^* + \text{herm.conj.}, \tag{4·13}$$

where

$$K = \begin{pmatrix} K^+ \\ K^0 \end{pmatrix} \quad \text{and} \quad \tau_2K^* = i\begin{pmatrix} -K^{0*} \\ K^{+*} \end{pmatrix} = i\begin{pmatrix} -\tilde{K}^0 \\ K^- \end{pmatrix},$$

is rearranged in the form

---

* Generally, the inversion of the $2n$-dimensional space can be represented by a rotation in the enlarged $(2n+1)$-dimensional space, as well known. The situation is very different, if the dimension of the original space is odd. But, by the analogy with the case of the even dimensional space, the inversion of the three-dimensional space can be written as the product of a rotation in the enlarged four-dimensional space $T_1T_2T_3T_4$ and the reflection with respect to the fourth axis T_4.

** It will be difficult to formulate the idea proposed in § 2 in such a way that we can construct the field which has the spinor character with respect to γ matrix and the tensor character with respect to ζ matrix. If we use Λ and Σ as it is, we must construct such fields.

$$\bar{N}YK^0+\bar{N}ZK^+ +\bar{\Xi}YK^- +\bar{\Xi}Z\bar{K}^0+\text{herm.conj.} \tag{4·14}$$

Namely, K meson cannot be arranged in the form of the spinor in isotopic spin space. But the form (4.14) suggests us to unify the strong interactions of K-meson into the form

$$\bar{\Psi}_{II}\,\mathcal{G}\,\Psi_{III}, \tag{4·15}$$

where

$$\begin{aligned}
&\Psi_{II}=\frac{1+\zeta_2\gamma_5}{2}\Psi_{II}+\frac{1-\zeta_2\gamma_5}{2}\Psi_{II}=\Psi_N+\Psi_\Xi\,,\\
&\Psi_{III}=\frac{1+\zeta_3\gamma_5}{2}\Psi_{III}+\frac{1-\zeta_3\gamma_5}{2}\Psi_{III}=\Psi_Y+\Psi_Z\,,
\end{aligned} \tag{4·16}$$

$$\begin{aligned}
\mathcal{G}=(\zeta_1-\zeta_3)\,\mathcal{G}'&=\frac{1+\zeta_2\gamma_5}{2}(\zeta_1-\zeta_3)\,\mathcal{G}'+\frac{1-\zeta_2\gamma_5}{2}(\zeta_1-\zeta_3)\,\mathcal{G}'\\
&=G_K+G_{\tilde{K}}\\
&=G_K\frac{1+\zeta_3\gamma_5}{2}+G_K\frac{1-\zeta_3\gamma_5}{2}\\
&\quad+G_{\tilde{K}}\frac{1+\zeta_3\gamma_5}{2}+G_{\tilde{K}}\frac{1-\zeta_3\gamma_5}{2}\\
&=G_{K^0}+G_{K^+}+G_{K^-}+G_{\tilde{K}^0}\,.
\end{aligned} \tag{4·17}$$

Of course, we cannot say anything at the present stage about the reason why the operators $\frac{1\pm\zeta_3\gamma_5}{2}$ pick up the various charge states of K-meson.*

The ζ-matrix is related to the Minkowski space through the parity mixture on the one hand, owing to the fact that it is contained in the wave equation. On the other hand, it is also connected to isotopic spin space by the equation (4·12). Therefore, if we enlarge the Minkowski space in such a way that the wave equation can be derived from the structure of the enlarged space as in the conventional theory, the enlarged space-time will not be orthogonal to the four-dimensional isotopic spin space. The non-orthogonality is due to the existence of the ζ matrix. Now, the appearance of the ζ-matrix is combined with the nonconservation of parity. Therefore, the nonorthogonality may be considered to be small, as the strength of the parity nonconserving interactions is small. Further, it will disappear in the limit where parity is conserved. Then, the orthogonality between the two spaces in the conventional theory is to be considered to have its meaning in the appproximation where parity is conserved.*

* The additional projection operator $(1\pm\zeta_3\gamma_5)/2$ does not disturb the correlation between two conservation laws of hyperonic charge and of parity.

The idea that the isotopic spin space has some internal correlation to the Minkowski space will change the conventional space-time concept radically, and seems to be too drastic. In the conventional theory, we start from the preliminarily arranged space-time, and fit the particle image to the framework. But the space-time structure is to be understood through physical phenomena. Namely, we must learn the particle image through the dynamics, and find some new structure of the space-time by means of the modification of the particle image**.

Of course, the required modification may not be such a simple one as to extend the four-dimensional Minkowski space to some six-dimensional space-time**, or to treat the isotopic spin space on the same foot as the Minkowski space. It will require some new type of field theory or mathematical tool to formulate the idea.

Appendix I

Here we want to emphasize that hyperonic charge is only one of the $2\times\infty$ possibilities of the phenomenological parameters which have the following properties, and that the essential point is the selection rule between the initial and final states. The selection rule is given by the following three requirements, as well known: a) the parameter is conserved through the production of baryons and mesons or their scattering. b) It changes by one unit in the decays of them. c) The reaction, in which it changes by more than two units, are forbidden.

At first, we may assign the parameter to baryons and mesons as follows:

$$a \text{ for } \pi,\quad b \text{ for } N,$$
$$c \text{ for } \theta^{\circ},\quad f \text{ for } \Lambda,$$
$$d \text{ for } K^{+},\ g \text{ for } \Sigma,$$
$$e \text{ for } K^{-},\ h \text{ for } \Xi.$$

Here we did not add $\tilde{\theta}^{\circ}$, because the production of it is not observed directly and in the decay it appears in the mode θ_1^0 or θ_2^0. If the production is observed, it can be treated in the same line as follows. Further, we grouped $\Sigma'^{\pm}$ and Σ'^{0} into one Σ', for simplicity, but it is not essential. In order to determine the values, it is sufficient to use the following nine events:

---

* Here we want to emphasize that the two spaces are not orthogonal in any sense, if we identify the ζ-matrix itself with the τ-matrix in the wave equations.

** The concepts of space-time and particle image are kinematical ones. In the conventional theory, the law, which prescribes the possible interactions (dynamics), is already comprised in the kinematical part of the theory. Namely, the Lagrangian of the free field is determined by the irreducible representation of the Lorentz group and the interactions are assumed to be invariant under the group. Therefore, if we keep such a ring, some modification of the dynamical part will induce the corresponding modification in the kinematical part.

*** A tentative model of such extension will be described in Appendix II.

production and scattering events

$$\pi^- + p \to \Lambda^0 + \theta^0,$$

$$\pi^- + p \to \Sigma^- + K^+,$$

$$K^- + p \to \Sigma^+ + \pi^-,$$

$$K^- + p \to \Lambda^0 + \pi^0,$$

decay events

$$\Lambda^0 \to \pi^- + p,$$

$$\Sigma^- \to p + \pi^0,$$

$$\theta^0 \to \pi^+ + \pi^-,$$

$$\Xi^- \to \Lambda^0 + \pi^-,$$

forbidden process

$$n + n \to \Lambda + \Lambda .$$

Then, after the elementary calculation, we get the following solutions,

$$\begin{aligned} &a=0,\ b=\alpha,\ c=d=-e=-1 , \\ &f=g=\alpha+1,\ h=(\alpha+1)\pm 1 , \end{aligned} \qquad \text{(A–I–1)}$$

and

$$\begin{aligned} &a=0,\ b=\alpha,\ c=d=-e=1 , \\ &f=g=\alpha-1,\ h=(\alpha-1)\pm 1 . \end{aligned} \qquad \text{(A–I–2)}$$

If we take into account the event[19]*

$$n + n \to \Xi^- + p + K^0 + K^0,$$

the value h becomes

$$h=\alpha+2 \qquad \text{(A–I–3)}$$

for the set (A–I–1), and

$$h=\alpha-2 \qquad \text{(A–I–4)}$$

for the set (A–I–2).

If we put $\alpha=0$, the parameter is equal to strangeness, and the two sets (A–I–1) and (A–I–2) are different only in the sign. If we put $\alpha=1$, the set (A–I–2) gives

* There are three possibilities about the two K-mesons appearing in the final state, that is (K^0, K^0), $(K^0, \tilde{K}^0)$ and $(\tilde{K}^0, \tilde{K}^0)$. But only (K^0, K^0) gives the consistent solution. We may assume that the change of the parameter is ± 2 in the forbidden process, and use the forbidden process $n+n \to \Lambda+\Lambda$, without referring to the above event. But the requirement (c) mentioned above is more general.

hyperonic charge.

Usually, $\tilde{\theta}^0$ was introduced in the following way. The charge and the concerned parameter of K^+ and K^- have the same magnitude and the inverse sign. Therefore, they are related to each other as particle and antiparticle. On the other hand, the properties of K^+ and θ^0 are very similar, apart from the charge. Then, K^+ and θ^0 were considered as the isotopic spin doublet. Consequently, θ^0 has its antiparticle state $\tilde{\theta}^0$, corresponding to the relation between K^+ and K^-. Therefore it appears to be essential that K-meson has the isotopic spin $\frac{1}{2}$. Indeed, the route of the consideration was very powerful, but, at the present stage, it may be necessary to reinvestigate the scheme of elementary particles, irrespective of the image. (See the discussion (D) in § 4, especially the equation (4.14).) Concerning this problem, the essential point will be to construct the theory in such a way that K^0 and K^+ have the same behaviour apart from the charge.

Appendix II.

We duplicated the number of the components of spinor, in order to interpret some phenomenological problems inclusively. Now, we must start from some six or seven-dimensional space in spinor analysis to treat the spinor whose number of components is eight.[20] It is very difficult to connect such an enlarged space to the Minkowski space. We know the theory of conformal transformation group[21] as the theory that makes a six-dimensional space correspond to the Minkowski space. But this theory includes some elements which make us hesitate to accept this group as a physical one.

In this appendix, we try to construct a six-dimensional space in such a way that we can get the equations (2.8–II) and (2.8–III). Then we assume for simplicity that the first four axises coincide with the ones in the Minkowski space. Of course, the future theory, which reproduces the idea explained in the text, will not have such a simple form. The following model is exhibited only for the purpose of facilitating the understanding of our idea, especially on the problems of the hyperonic charge of bosons and of interactions.

We associate the hermitian matrix H_i, which satisfies the commutation relation

$$\{H_i,\ H_j\}=2\delta_{ij}\ \ (i,j=1,2,\cdots 6), \qquad \text{(A–II–1)}$$

with the unit vector $x^{(i)}$ in the direction of the i-th axis in our space. Then we can associate the matrix

$$X=x_iH_i$$

with an arbitrary vector x, and we have

$$(x_i\ y_i)=\tfrac{1}{2}(XY+YX).$$

If we consider a reflection with respect to a plane whose unit normal vector is given by n_i, the vector matrix X is transformed into

$$X'=-NXN^{-1},$$

and we can define the spinor χ as the quantity which is transformed into

$$\chi' = N\chi$$

under the reflection. The arbitrary transformation in our six-dimensional space can be decomposed into the product of some reflections, as well known. If we express the ρ-th reflection operator by $N^{(\rho)}$, the concerned transformation is represented as

$$\chi' = S\chi ,$$

where

$$S = N^{(\nu)} N^{(\nu-1)} \cdots N^{(2)} N^{(1)} . \tag{A–II–3}$$

Then, the vector matrix X is transformed into

$$X' = \sigma S X S^{-1} \tag{A–II–4}$$

under the transformation,* and we obtain the relation

$$H_i a_{ij} = \sigma S H_j S^{-1}, \tag{A–II–5}$$

for the coordinate transformation

$$x_i' = a_{ik}\ x_k. \tag{A–II–6}$$

The inverse operator of S can be rewritten as

$$S^{-1} = \sigma_t N^{(1)} N^{(2)} \cdots N^{(\nu)}, \tag{A–II–7}$$

if we take into account

$$N^2 = (n_i n_i) = \pm 1, \tag{A–II–8}$$

according as the direction of the vector n is space-like or time-like. Next, we define the spinor $\tilde{\chi}$ associated with χ as

$$\tilde{\chi} = \chi B \tag{A–II–9}$$

by means of the matrix B, which satisfies the relation

$$N^{\dagger} B = -BN \tag{A–II–10}$$

for an arbitrary reflection matrix N. Then we can derive the relation

$$S^{\dagger} B = \sigma_s B S^{-1} \tag{A–II–11}$$

at once. Therefore, the spinor $\tilde{\chi}$ is transformed into

$$\tilde{\chi}' = \sigma \tilde{\chi} S^{-1}, \tag{A–II–12}$$

* σ, σ_t and σ_s are the sign functions defined as

$$\sigma = (-1)^{\nu},\ \sigma_t = (-1)^{\nu_t} \text{ and } \sigma_s = (-1)^{\nu_s},$$

where ν, ν_t and ν_s are the numbers of the total, time-like-and space-like-reflections involved in S.

under the transformation S. The matrix B depends on the metric in our space, and we may suitably assume the fundamental form as

$$\sum_{i=1}^{6} x_i^2,$$

supposing x_4 to be pure imaginary and all other coordinates to be real. Then we have

$$B=H_4. \tag{A–II–13}$$

Here we introduce the Γ-system defined by

$$\Gamma_i=iH_7H_i \text{ and } \Gamma_7=H_7, \tag{A–II–14}$$

where

$$H_7=iH_1H_2\cdots H_6. \tag{A–II–15}$$

As to the Γ-system, we have the relations

$$\Gamma_i a_{ij}=S\Gamma_j S^{-1} \tag{A–II–5′}$$

and

$$S^{\dagger}\Gamma_4=\sigma_t\,\Gamma_4 S^{-1} \tag{A–II–11′}$$

instead of (A–II–5) and (A–II–11), respectively. Therefore, the adjoint spinor $\bar{\chi}$, defined by

$$\bar{\chi}=\chi^{\dagger}\,\Gamma_4, \tag{A–II–9′}$$

is transformed into

$$\bar{\chi}'=\sigma_t\bar{\chi}S^{-1}, \tag{A–II–12′}$$

and $\bar{\chi}\chi$ becomes the scalar quantity, except for the factor σ_t. We can take

$$\mathcal{L}=\bar{\chi}(\Gamma_i\,\partial_i+k)\chi \tag{A–II–16}$$

as the most general invariance, (except for the factor σ_t), from which we can derive the wave equation.

In the following, we fix the representation of Γ_i's as*

$$\begin{aligned}\Gamma_\mu&=i\zeta_1\gamma_5\gamma_\mu,\\ \Gamma_5&=\zeta_3,\\ \Gamma_6&=\zeta_2,\end{aligned} \tag{A–II–17}$$

and

in order to derive the equations which are introduced in the text. Then the eigenvalue equations for the matrices $\zeta_i\gamma_5$'s, the equations (2.10), are written down in terms of the

* In the spinor equation, the representation is of no importance, as far as we are concerned with this six-dimensional space. The speciality of some ζ matrix appears in connection with such extended isotopic spin space as given in the equation (4.14).

representation as

$$\Gamma_7\chi_I^{\pm}=\pm\chi_I^{\pm}, \qquad \text{(A–II–18–I)}$$

$$i\Gamma_7\Gamma_5\chi_{II}^{\pm}=\pm\chi_{II}^{\pm} \qquad \text{(A–II–18–II)}$$

and

$$i\Gamma_6\Gamma_7\chi_{III}^{\pm}=\pm\chi_{III}^{\pm}, \qquad \text{(A–II–18–III)}$$

respectively.

The Lagrangian density, $\mathcal{L}_{II}$, which gives the wave equation for Ψ_{II}, (2·8–II), must be invariant for the substitution (A–II–18–II). The term involving Γ_5 does not satisfy this invariance character. Further, the k-term has the same character as the Γ_5 term in our equation, as will be seen in the following. Therefore, we may put

$$\mathcal{L}_{II}=\bar{\chi}_{II}(\Gamma_\mu\partial_\mu-\Gamma_5\partial_5)\chi_{II}. \qquad \text{(A–II–19–II)}$$

The wave equation

$$(\Gamma_\mu\partial_\mu+\Gamma_6\partial_6)\chi_{II}=0 \qquad \text{(A–II–20–II)}$$

can be transformed into

$$(e^{2mx_6\Gamma_6}\ \Gamma_\mu\ \partial_\mu+m)\,\Psi_{II}=0\,, \qquad \text{(A–II–21–II)}$$

if we assume the x_6-dependence of χ_{II} as

$$\chi_{II}=e^{-mx_6\Gamma_6}\ \Psi_{II}\,. \qquad \text{(A–II–22–II)}$$

The equation (A–II–21–II) can be written, in terms of the representation (A–II–17), as

$$(i(\alpha+i\beta\zeta_2^z)\zeta_1^y\gamma_5\gamma_\mu\partial_\mu+m)\,\Psi_{II}=0. \qquad \text{(A–II–23–II)}$$

This equation coincides with the equation (2.8–II).*

The corresponding equations for χ_{III} are obtained similarly as

$$\mathcal{L}_{III}=\bar{\chi}_{III}(\Gamma_\mu\partial_\mu+\Gamma_5\partial_5)\chi_{III}, \qquad \text{(A–II–19–III)}$$

$$(\Gamma_\mu\partial_\mu+\Gamma_5\partial_5)\chi_{III}=0, \qquad \text{(A–II–20–III)}$$

$$(e^{-2imx_5\Gamma_5}\ \Gamma_\mu\partial_\mu+im)\,\Psi_{III}=0, \qquad \text{(A–II–21–III)}$$

$$\chi_{III}=e^{imx_5\Gamma_5}\ \Psi_{III} \qquad \text{(A–II–22–III)}$$

and

$$\{(a+b\zeta_3^z)\zeta_1^y\gamma_5\gamma_\mu\partial_\mu+m\}\,\Psi_{III}=0. \qquad \text{(A–II–23–III)}$$

* We can consider x_6 as the parameter, which represents the order of the parity mixing, because the limit $x_6\to 0$ corresponds to the limit $\alpha\to 1$ and $\beta\to 0$. The situation that the transformation (A-II-22-II) is not unitary may be considered as unsatisfactory, but it corresponds to the same situation for the transformation (2.5). The equation (A-II-21-II) can also be interpreted as the one for a parity doublet state, if we rewrite the equation as

$$(\Gamma_\mu\partial_\mu+me^{-2mx_6\Gamma_6})\,\Psi_{II}=0. \qquad \text{(A-II-21′-II)}$$

The last equation has the same form as (2.8–III).

The substitution (A–II–18–I) keeps the Lagrangian (A–II–13) invariant, except for the k-term, but it is impossible to derive the equation (2·8–I) for Ψ_I from this Lagrangian*. We need not labour at the formulation which derives all types of spinor fields simultaneously, as we can analyse the concerned phenomena without referring to Ψ_I. Further, it may be interesting that we can write (A–II–22–II) and (A–II–22–III) as

$$\chi = e^{im\varepsilon\Gamma_5\Gamma_6\Gamma_7}\Psi = e^{m\varepsilon\Gamma_5}\Psi, \tag{A–II–24}$$

where

$$x = ix_6 \quad \text{for} \quad \chi_{II^+}$$

and

$$x = x_5 \quad \text{for} \quad \chi_{III^+},$$

if we take (A–II–18) into account.

The particle-antiparticle conjugation operator in our space can be determined so as to satisfy the relation

$$\begin{aligned}&\bar{\chi}^c(\Gamma_\mu\,\partial_\mu+\Gamma_5\,\partial_5+\Gamma_6\,\partial_6+k)\chi^c\\&\quad=\chi^T(\Gamma'^T_\mu\,\partial_\mu+\Gamma'^T_5\,\partial_5+\Gamma'^T_6\,\partial_6-k)\bar{\chi}^T,\end{aligned} \tag{A–II–25}$$

or

$$\bar{\chi}^c\,\Gamma_i\,\chi^c=\chi^T\,\Gamma'^T_i\,\bar{\chi}^T$$

and

$$\bar{\chi}^c\,\chi^c=-\chi^T\,\bar{\chi}^T. \tag{A–II–26}$$

If we put

$$\chi^c=C\,\bar{\chi}^T, \tag{A–II–27}$$

the requirements (A–II–26) impose the relations

$$\begin{aligned}&C\,C=1\\&\Gamma'^T_i=-C^{-1}\,\Gamma_i\,C\end{aligned} \tag{A–II–28}$$

on the operator C. In our representation (A–II–16), these relations determine the operator C as

$$C=\zeta_2\,\rho_1\,\sigma_2. \tag{A–II–29}$$

In the text we used the operator A defined by $\Psi^c=A\Psi$, and this operator can be expressed as

* We can derive the wave equation (2.8-I), if we put $\chi=e^{im\varepsilon\Gamma_5\Gamma_6}\Psi$ in the equation (A-II-20-II). Therefore, it may be permissible to consider that the equation (A-II-20-II) contains such possibility.

$$A = C\,\Gamma_4^T\,K$$
$$= \zeta_3\,\zeta_3\,\sigma_2\,K. \qquad \text{(A–II–30)}$$

This operator coincides with the operator A_{II} given in the equation (2·11–II). Further, our definition of the adjoint spinor, (A–II–9′), is just the same one as the definition (2.15–II), if we take into account that $\bar{\Psi}_{II} = \Psi_{II}^{\dagger}\,\gamma_4\,\zeta_3$ is equivalent to $\Psi_{II}^{\dagger}\,\zeta_2\,\gamma_5\cdot\gamma_4\zeta_3 = i\Psi_{II}^{\dagger}\,\zeta_1\,\gamma_5\,\gamma_4$.

We may consider that our six-dimensional space is the suitable one for the particle image of the field Ψ_{II}, as the concepts of adjoint and conjugate fields are determined by the structure of the fundamental space. As to Ψ_{III}, our space can introduce the wave equation for Ψ_{III}, equation (A–II–21 III), but $\bar{\Psi}_{III}$ and $\Psi_{III}^{\dagger}$ are different from the ones defined in the text.* Further, it may be considered as a reflection of this situation that the transformation which connects Ψ_{III} to χ_{III} is unitary, contrary to the transformation (A–II–22 -II). But such space is sufficient for our purpose, as the essential point is the operator $\zeta_2\gamma_5$ which characterizes the spinor Ψ_{II}.

Thus, the idea explained in the text can be represented by introducing the six-dimensional space whose fundamental reflection operators are given by (A–II–17) and (A–II–14), as long as we are concerned with spinor fields. But we cannot accept this space as the physical one as it is, because the meaning of the x_5 and x_6 coordinates, which are added to the Minkowski space, is not clear.**

In the conventional field theory, the structure of the space is represented by the spinor representation, and the properties of boson fields are derived from the representation by means of the so-called fusion technique[13]. In our scheme, the space is only a temporal model, and it will be questionable to extrapolate the technique to our space. Nevertheless, we are obliged to use the method with a slight modification, as there is no other way to extend the theory to boson fields at the present stage.

We introduce a tensor G which has the same transformation character as $\chi\chi'$. Then G is transformed into

$$G' = \sigma_t S G S^{-1}, \qquad \text{(A–II–31)}$$

as will be easily seen from the equations (A–II–2) and (A–II–12′). This tensor can be expanded by means of the 64 independent matrices as***

$$G = A + A'\Gamma_7 + A_i\,\Gamma_i + iA_i'\,\Gamma_i\,\Gamma_7$$
$$+ iA_{ij}\,\Gamma_{ij} + iA'_{ij}\,\Gamma_{ij}\,\Gamma_7 + iA_{ijk}\,\Gamma_{ijk}, \qquad \text{(A–II–32)}$$

where

$$\Gamma'_{ij} = \frac{1}{2}(\Gamma'_i\,\Gamma'_j - \Gamma'_j\,\Gamma'_i)$$

* The space which is suitable to Ψ_{III} is obtained if we assume x_5 to be imaginary in addition to x_4.

** The magnitudes of x_5 and x_6 have some physical meaning, as mentioned above. But, the meaning of the rotation in the plane, which includes x_5 or x_6, is quite obscure.

*** The factor i is added in such a way that G is hermitian when A...'s are real.

and

$$\Gamma_{ijk}=\tfrac{1}{6}(\Gamma_i\Gamma_j\Gamma_k-\Gamma_i\Gamma_k\Gamma_j+\Gamma_k\Gamma_i\Gamma_j \\ -\Gamma_k\Gamma_j\Gamma_i+\Gamma_j\Gamma_k\Gamma_i-\Gamma_j\Gamma_i\Gamma_k).$$

It can be easily shown that the coefficients A, A', A_i, A_i', A_{ij}, A_{ij}', and A_{ijk} have the character of scalar, pseudoscalar, vector, pseudovector, second rank tensor, second rank pseudotensor, and third rank tensor with respect to the transformation in our six-dimensional space, respectively, as will be proved by means of the properties (A–II–31) and (A–II–5′).

The necessary field quantities are the ones which have the pseudoscalar character with respect to the transformation in the Minkowski space embedded in the six-dimensional space. If we pick up such world-pseudoscalar, they are given as*)

$$A'\Gamma_7+A_5\Gamma_5+A_6\Gamma_6+iA'_{56}\Gamma_5\Gamma_6\Gamma_7. \qquad \text{(A–II–33)}$$

Here we rewrite the pseudoscalar part of G in terms of ζ and γ matrices as

$$G=A'\zeta_1\gamma_5+A_5\zeta_3+A_6\zeta_2+A'_{56}\gamma_5. \qquad \text{(A–II–34)}$$

In order to introduce hyperonic charge, we construct the tensor matrices $G_{II}\pm$ as

$$G_{II}\pm=\frac{1\pm\zeta_2\gamma_5}{2}G$$

$$=\frac{1\pm\zeta_2\gamma_5}{2}\{(A_5\mp iA')\zeta_3+(A_{56}\pm A_6)\gamma_5\}. \qquad \text{(A–II–35)}$$

The tensor matrices $G_{II}^{(1)}\pm$ and $G_{II}^{(2)}\pm$ defined by

$$G_{II}^{(1)}\pm=(A_5\mp iA')\frac{1\pm\zeta_2\gamma_5}{2}\zeta_3$$

and $\qquad \text{(A–II–36)}$

$$G_{II}^{(2)}\pm=(A'_{56}\pm A_6)\frac{1\pm\zeta_2\gamma_5}{2}\gamma_5$$

have the properties

$$\zeta_2\gamma_5G_{II}^{(1)}\pm=-G_{II}^{(1)}\pm\zeta_2\gamma_5=\pm G_{II}^{(1)}\pm$$

and $\qquad \text{(A–II–37)}$

$$\zeta_2\gamma_5G_{II}^{(2)}\pm=G_{II}^{(2)}\pm\zeta_2\gamma_5=\pm G_{II}^{(2)}\pm,$$

respectively. If we compare these properties with the equations (2.23∼25), we may try to identify G_π, G_K and $G_{\bar{K}}$ with $G_{II}^{(2)}+$, $G_{II}^{(1)}+$ and $G_{II}^{(1)}-$, respectively. The only difference between them is the γ_5-dependence of G_K and $G_{\bar{K}}$. Namely, G_K can be rewrit-

* For instance, the term $A_6\Gamma_6$ transforms under the Lorentz transformation, in which the matrix S involves only Γ_μ (μ=1,2,3,4), as

$$A_6\Gamma_6\to\sigma_\ell S(A_6\Gamma_6)S^{-1}=\sigma_\ell A_6\Gamma_6.$$

ten as $\frac{1+\zeta_2\gamma_5}{2}(\zeta_1-\zeta_3)=\frac{1+\zeta_2\gamma_5}{2}(-i\gamma_5-1)\zeta_3$, contrary to $G_{II}^{(1)+}$. Here we study the properties of these tensors for the particle-antiparticle conjugation. Taking into account the transformation (A–II 27) and the definition that G transforms as $\chi\chi'$, G is transformed into

$$-C\Gamma_4^T G^* \Gamma^T{}_4 C^{-1} \tag{A–II–38}$$

under this transformation. Therefore, $G^{(1)}$ and $G^{(2)}$ are transformed into

$$G_{II}^{(1)\pm}\rightarrow -(A_5\pm iA')^*\frac{1\mp\zeta_2\gamma_5}{2}\zeta_3,$$

and (A–II–39)

$$G_{II}^{(2)\pm}\rightarrow -(A'_{5\mu}+A_\mu)^*\frac{1\mp\zeta_2\gamma_5}{2}\gamma_5,$$

respectively. Namely $G_{II}^{(1)+}$ is converted into $G_{II}^{(2)-}$, and this character corresponds to the particle-antiparticle relation between K- and $\tilde{K}$-mesons, as will be expected. But, $G_{II}^{(2)+}$, which may be identified with G_π, is transformed into $G_{II}^{(2)-}$. This situation and the unsuitable γ_5-dependence of G_K show that the use of the fusion technique in such a simple manner is unsatisfactory for our scheme.

The Yukawa type interaction is expressed as

$$\mathrm{sp}(G\,\Psi\,\bar{\Phi}), \tag{A–II–40}$$

where G is given by (A–II–34). This expression is a scalar in the six-dimensional space, as will be easily seen. In order to show the way in which the possible interactions are picked up, we take as an example the case where $\Psi=\Psi_{III}$ and $\Phi=\Psi_{II}$. Then

$$\begin{aligned}&\mathrm{sp}(G\,\Psi_{III}\bar{\Psi}_{II^+})\\&=\mathrm{sp}\left(G\,\Psi_{III}\bar{\Psi}_{II^+}\frac{1+i\Gamma_7\Gamma_5}{2}\right)\\&=\mathrm{sp}\left(\frac{1+i\Gamma_7\Gamma_5}{2}G\,\Psi_{III}\bar{\Psi}_{II^+}\right)\\&=\mathrm{sp}\{(G_{II}^{(1)+}+G_{II}^{(2)+})\,\Psi_{III}\bar{\Psi}_{II^+}\}.\end{aligned} \tag{A–II–41}$$

Therefore, there are $\bar{N}\Lambda K$, $\bar{N}\Sigma K$, $\bar{N}\Lambda\pi$ and $\bar{N}\Sigma\pi$ interactions.

The total Lagrangian is invariant under the transformation

$$\chi\rightarrow\Gamma_7\chi, \tag{A–II–42}$$

but the meaning of this transformation is not clear at the present stage.

Before concluding this appendix, we want to make a speculation to connect the Minkowski space to the isotopic spin space. In our representation, the reflection operators H_5 and H_6 are given by $\zeta_2\gamma_5$ and $-\zeta_3\gamma_5$ respectively, but we can reconstruct our scheme

in such a way that $H_5=\zeta_3$ and $H_6=\zeta_2$. If we put $\zeta_j=\zeta_3$ in the equation (4·12), we have $T_4=H_5$, and this can be understood as to show that the x_5 axis has the same role as the fourth axis in the extended isotopic spin space, with respect to the reflection. In our scheme the baryon fields have $8\times 2=16$ components. On the other hand, we have six coordinates x_i in the extended Minkowski space, and four coordinates t_μ in the extended isotopic spin space. If we identify t_4 with x_5, there remain nine coordinates. Generally, the spinor in the nine-dimensional space has sixteen components, and this number coincides with the number of the components of our spinor. Of course, the connection between the two spaces is not such a docile one as the simple nine-dimensional space, and some new mathematics will be necessary to treat such connection, even if such classical connection has any physical meaning. The correlation between two spaces discussed at the end of § 4 may be described only by the quantum theoretical method, for example, through the structure of Hilbert space.

References

1) M. Gell-Mann Phys. Rev. **92** (1953), 833.
T. Nakano and K. Nishijima Prog. Theor. Phys. **10** (1953), 581.
K. Nishijima Prog. Theor. Phys. **12** (1954), 107 and **13** (1955), 285.
2) C. S. Wu et al. Phys. Rev. **105** (1957), 1413.
R. L. Garwin et al. Phys. Rev. **105** (1957), 1415.
3) F. Eisler et al. Phys. Rev. **108** (1958), 1353.
4) T. D. Lee and C.N. Yang Phys. Rev. **102** (1956), 290.
5) M. Gell-Mann Phys. Rev. **97** (1955), 1387.
6) M. Gell-Mann and A. Pais Phys. Rev. **106** (1957), 1296.
7) S. Sakata Prog. Theor. Phys. **16** (1956), 686
S. Tanaka Prog. Theor. Phys. **16** (1956), 625, 631.
8) B. d'Espagnat and J. Prentki Nuclear Physics **1** (1956), 33.
9) G. Racah Nuclear Physics **1** (1956), 657
Y. Murai Nuclear Physics **1** (1956), 302.
B. d'Espagnat and J. Prentki Phys. Rev. **102** (1956), 1684.
10) J. Schwinger Phys. Rev. **104** (1956), 1164.
11) K. Lande et al. Phys. Rev. **105** (1957), 1925.
W. K. H. Panofsky et. al. Phys. Rev. **109** (1958), 1353
12) T. D. Lee and C. N. Yang Phys. Rev. **104** (1956), 254.
13) See, for example, S. Watanabe Phys. Rev. **106** (1957), 1306.
14) T. D. Lee and C. N. Yang Phys. Rev. **105** (1957), 1671
L. Landau Nuclear Physics **3** (1957), 127.
A. Salam Nuovo Cimento **5** (1957), 299.
15) E. J. Konopinski and H. M. Mahmoud Phys. Rev. **92** (1953), 1045.
16) See for example, K. Huang and F. E. Low Phys. Rev. **109** (1958), 1400.
17) T. Okabayashi Prog. Theor. Phys. **10** (1953), 499.
18) See for example M. Kawaguchi Phys. Rev. **107** (1957), 573.
R. H. Dalitz Proc. Phys. Soc. **A69** (1956), 527.
19) C. D. Anderson et al. Phys. Rev. **100** (3955), 1457.
20) See, for example E. Cartan *Leçons sur la théorie des spineurs* (1938) Hermann, Paris.
21) See, for example P.A.M. Dirac Annals of Mathematics **37** (1936), 429
Y. Murai Prog. Theor. Phys. **11** (1954), 441.
F. Klein *Vorlesungen über höherere Geometrie* (1926), Springer p. 246
22) S. Watanabe Phys. Rev. **84** (1951), 1008, Appendix.

Progress of Theoretical Physics, Vol. 20, No. 5, November 1958

Bose-Einstein Lattice Gases equivalent to the Heisenberg Model of Ferro-, Antiferro- and Ferri-Magnetism

Tohru MORITA

Physics Department, Tokyo Institute of Technology, Tokyo

(Received June 30, 1958)

The Hamiltonians are presented which are equivalent to the Heisenberg model of ferro-, antiferro- and ferri-magnetism and which have the form of finite power series of the Bose operators.

A technique for calculations of partition function for a system with such a Hamiltonian that contains a term of infinite potential is presented and discussed in connection with the theories of a ferromagnetic spin system by Van Kranendonk and Dyson.

§ 1. Introduction*

Bloch[2] introduced the spin waves as the lowest excitations of the Heisenberg model of ferromagnetics and succeeded in explaining the $T^{3/2}$ law of the variation of the spontaneous magnetization, by neglecting the interactions of the spin waves. The spin wave introduced by Bloch follows the law of Bosons, while the spin reversal operators $S_j^{\pm}=S_{jx}\pm iS_{jy}$ at a lattice site j follow that of Fermions. Hence, to get the spin wave picture, we must write $S_j^{\pm}$ in terms of some operators of the Bose nature, a_j, a_j^*. This was done by Holstein and Primakoff[3]. However, this was done in such a form as to express $S_j^{\pm}$ in an infinite power series of a_j and a_j^*, though in a simple analytic form. When we start with their Hamiltonian in terms of a_j and a_j^*, we are forced to cut the infinite series retaining only a few terms with small powers. In such an approximation, it is pointed out that many divergence difficulties arise and they are attributed to the spreading out of the Hilbert space caused by the neglect of the higher terms.[4] On the other hand, the spin wave formalism has been applied to the antiferromagnetics by Anderson[1] and shown to lead to a good estimation of the antiferromagnetic ground state energy. To this case Holstein and Primakoff's way to write the Hamiltonian in terms of the Bose operators is applied by Kubo.[4] In this case also, difficulties arise, which are attributed to the spreading out of the Hilbert space. It will be impossible, however, to cope with these difficulties by considering the infinite series in full terms.

It is the purpose of this paper to present a Hamiltonian that has the form of a finite series of Bose operators and is equivalent to the Heisenberg model of ferro-, antiferro- and ferri-magnetism.** This can be done very easily when we introduce an infinite

* Quite recently a review article has been published by Van Kranendonk and Van Vleck[5] in which the spin wave theory is discussed without including their interactions.

** The purpose of this paper is to present a definite basis on which the theory of magnitism can be constructed. The construction of the theory on this basis will be a subject of future investigations.

potential to prevent the spreading out of the Hilbert space. This is the correspondence, for the cases of spin $S \geq 1/2$, of the proof due to the author[6] that the eigenvalue problem of a Hamiltonian of a Bose lattice gas with hard core is equivalent to that of the Hamiltonian of an anisotropic spin system with spin $1/2$. The latter fact has been proved in a more elementary but less general way by Van Kranendonk[9] in the reverse direction to the author's. Then Van Kranendonk,[9)10)] using this picture calculated the characters of ferromagnetics by treating the equivalent Bose lattice gas instead of the ferromagnetics directly. This is just the reverse of the treatment of Matsubara and Matsuda[7,8)] who treated the equivalent spin system to investigate the properties of a Bose lattice gas.

The Hamiltonian to be obtained contains an infinite potential, accordingly in our development the difficulties due to the spreading out of the Hilbert space may well be said to be replaced by those due to the infinite potential. Even with them, the Hamiltonian to be obtained has a feasible character in the physical point of view because of the form of finite power series.

In §§ 2, 3 and 4, are obtained the Hamiltonians equivalent to those of ferro-, antiferro- and ferri-magnetics, respectively, with anisotropy and a magnetic field parallel to the axis of anisotropy.

At last, in § 5, we discuss a technique available for the calculation of the partition function for a system with such a Hamiltonian that contains a term of infinite potential which restricts the Hilbert space. In connection with this technique, the theories of Van Kranendonk[9)10)] and Dyson[11)12)] are compared with.

§ 2. Ferromagnetism

In this section, we consider the Hamiltonian of ferromagnetics of spin S, each on L lattice sites, with an anisotropy and a magnetic field H parallel to the axis of anisotropy; which is chosen to be z axis:

$$\begin{aligned}\mathcal{H}_0 &= -2J\sum_{\langle jk\rangle}{}'(S_{jx}S_{kx}+S_{jy}S_{ky}) - 2J'\sum_{\langle jk\rangle}S_{jz}S_{kz} - H\sum_{j=1}^{L}S_{jz} \\ &= -J\sum_{\langle jk\rangle}{}'(S_j^-S_k^+ + S_j^+S_k^-) - 2J'\sum_{\langle jk\rangle}{}'S_{jz}S_{kz} - H\sum_{j=1}^{L}S_{jz}\,, \qquad (2\cdot1)\end{aligned}$$

$$S_j^{\pm} = S_{jx} \pm iS_{jy}\,, \qquad (2\cdot2)$$

where J and J' are exchange integrals in x and y directions and z direction respectively, S_x, S_y and S_z are well-known angular momentum operators in the unit of $\hbar$, and $\langle jk\rangle$ means the summation over the nearest neighbor lattice sites pairs. Because of their definition, $S_j^{\pm}$ satisfy the commutation relations:

$$\begin{aligned}&S_j^-S_j^+ + S_j^+S_j^- = 2S(S+1) - 2S_{jz}^2\,, \\ &[S_j^-,\ S_k^+] = -2\delta_{j,k}S_{jz}, \qquad [S_j^{\pm},\ S_k^{\pm}] = 0.\end{aligned} \qquad (2\cdot3)$$

That is to say, S_j^-'s and S_k^+'s are commutative for different lattice sites, while they are not so

for a lattice site. In order to write the Hamiltonian in terms of the spin wave operators, we must first write the Hamiltonian in terms of the Bose operators, a_j and a_j^*, which are commutative also for a lattice site:

$$[a_j, a_k^*]=\delta_{j,k}, \qquad [a_j, a_k]=0. \tag{2·4}$$

This was done by Holstein and Primakoff[3] by noticing that S_j^-, S_j^- and S_{jz} in the representation where S_{jz} is diagonal, are equal to the parts $0\leq n_j=a_j^*a_j\leq 2S$ of

$$\begin{aligned}&(2S)^{1/2}(1-a_j^*a_j/2S)^{1/2}a_j,\\ &(2S)^{1/2}a_j^*(1-a_j^*a_j/2S)^{1/2}\\ &S-a_j^*a_j\end{aligned} \tag{2·5}$$

respectively, in the representation where $a_j^*a_j$ is diagonal, and that, when these are used in place of $S_j^{\pm}$ and S_{jz} in the Hamiltonian (2·1), the Hilbert space of eigensolutions is seperated into the regions where $a_j^*a_j\leq 2S$ and $a_j^*a_j>2S$.[13] That is, our problem reduces to the eigenvalue problem of

$$\begin{aligned}\mathcal{H}_0=&-LzS^2J'-LSH+(2zSJ'+H)\sum_{j=1}^{L}a_j^*a_j-2J'\sum_{<jk>}a_j^*a_ja_k^*a_k\\ &-2SJ\sum_{<jk>}\left\{a_j^*\left(1-\frac{a_j^*a_j}{2S}\right)^{1/2}\left(1-\frac{a_k^*a_k}{2S}\right)^{1/2}a_k+a_k\left(1-\frac{a_k^*a_k}{2S}\right)^{1/2}\left(1-\frac{a_j^*a_j}{2S}\right)^{1/2}a_j^*\right\}\end{aligned} \tag{2·6}$$

in the Hilbert space*

$$\Pi_{j=1}^{L}(a_j^*)^{n_j}\Phi_0 \qquad (n_j=0, 1, \cdots, 2S), \tag{2·7}$$

where Φ_0 is the state satisfying

$$a_j\Phi_0=0 \qquad \text{for all } j$$

and $(\Phi_0, \Phi_0)\neq 0$; z is the coordination number of the lattice. The spin wave theory of Holstein and Primakoff[3] is to develop $(1-a_j^*a_j/2S)^{1/2}$ in powers of $a_j^*a_j$ in (2·6) and neglect the higher powers of $a_j^*a_j$ than the first order and to solve the eigenvalue problem in the Hilbert space

$$\Pi_{j=1}^{L}(a_j^*)^{n_j}\Phi_0 \qquad (n_j=0, 1, \cdots, \infty), \tag{2·8}$$

neglecting the restriction of $n_j\leq 2S$ (i.e. neglecting the so-called "kinematical interactions").

Now, we expand $(1-a_j^*a_j/2S)^{1/2}$ as**

---

* This Hilbert space is identical to the original one:

$$\Pi_{j=1}^{L}(S_j^-)^{n_j}\Phi_0 \qquad (n_j=0, 1, \cdots, 2S) \tag{2·7'}$$

where Φ_0 is the state characterized by

$$S_j^+\Phi_0=0 \qquad \text{for all } j$$

and $(\Phi_0, \Phi_0)\neq 0$.

** The author was remarked that this expansion has already been used by T. Kasuya in *Busseiron-Kenkyu* **92** (1956), 14.

$$\left(1-\frac{a_j^* a_j}{2S}\right)^{1/2}=\sum_{n=0}^{\infty}\frac{c_n}{n!}a_j^{*n}a_j^n\ , \tag{2·9}$$

then the parts $0\leqq n_j=a_j^* a_j\leqq 2S$ of

$$\begin{aligned}&(2S)^{1/2}\sum_{n=0}^{2S-1}\frac{c_n}{n!}a_j^{*n}a_j^n a_j\\ &(2S)^{1/2}a_j^*\sum_{n=0}^{2S-1}\frac{c_n}{n!}a_j^{*n}a_j^n\end{aligned} \tag{2·10}$$

in the representation where $a_j^* a_j$ is diagonal, are also equal to S_j^+ and S_j^- in the representation where S_{jz} is diagonal. So that, our problem can be taken also as the eigenvalue problem of

$$\mathcal{H}_0=-LzS^2J'-LHS+(2zSJ'+H)\sum_j a_j^* a_j-2J'\sum_{<jk>}a_j^* a_j a_k^* a_k$$

$$-2SJ\sum_{<jk>}\left\{a_j^*\sum_{n=0}^{2S-1}\frac{c_n}{n!}a_j^{*n}a_j^n\sum_{m=0}^{2S-1}\frac{c_m}{m!}a_k^{*m}a_k^m a_k+a_k^*\sum_{m=0}^{2S-1}\frac{c_m}{m!}a_k^{*m}a_k^m\sum_{n=1}^{2S-1}\frac{c_n}{n!}a_j^{*n}a_j^n a_j\right\} \tag{2·11}$$

in the Hilbert space (2·7). The coefficients c_n's are calculated in Appendix.

Our next step is to remove the cumbersome restriction $n_j\leqq 2S$. For this, we have only to add to our Hamiltonian (2·11) a term which is infinite when the state is outside the original Hilbert space (2·7) and is zero when the state is inside it. Even when such a term is added to the Hamiltonian (2·11) and its eigenvalue problem in the Hilbert space (2.8) is solved, the set of eigenvalues and eigenfunctions is obviously unchanged, except for the appearance of states with infinite eigenvalue. As such a term we can take

$$\mathcal{H}'=v_0\sum_{j=1}^{N}a_j^{*2S+1}a_j^{2S+1}\ ,\qquad v_0=\infty\ . \tag{2·12}$$

As the result, we get the Hamiltonian

$$\begin{aligned}\mathcal{H}&=\mathcal{H}_0+\mathcal{H}'\\ &=\text{Eq. (2·11)}+\text{Eq. (2·12)}\end{aligned} \tag{2·13}$$

in the Hilbert space (2·8) as an equivalent to the Hamiltonian (2·1) in the Hilbert space (2·7′).

It is noticed that this Hamiltonian commutes whith $\sum_{j=1}^{N}a_j^* a_j$. Then, this Hamiltonian is a Hamiltonian of $N(=\sum_{j=1}^{N}a_j^* a_j)$ interacting Bose particles. The presence of $v_0=\infty$ introduces the difficulties analogous to those met in the theory of liquid or gas which is caused by the divergence of the Fourier transform of the potential of interaction.

Using the result of Appendix for the values of c_n's, we write down this Hamiltonian explicitly for $S=1/2$, 1 and $3/2$ in the following:

For $S=1/2$*

$$\mathcal{H}=-\frac{Lz}{4}J'-\frac{1}{2}LH+(zJ'+H)\sum_{j=1}^{L}a_j^*a_j-2J'\sum_{<jk>}a_j^*a_ja_k^*a_k$$
$$-J\sum_{<jk>}(a_j^*a_k+a_k^*a_j)+v_0\sum_{j=1}^{L}a_j^*a_j^*a_ja_j\,,\qquad v_0=\infty. \tag{2·14}$$

For $S=1$

$$\mathcal{H}=-LzJ'-LH+(2zJ'+H)\sum_{j=1}^{L}a_j^*a_j-2J'\sum_{<jk>}a_j^*a_ja_k^*a_k$$
$$-2J\sum_{<jk>}\Big\{a_j^*\Big(1-\frac{\sqrt{2}-1}{\sqrt{2}}a_j^*a_j\Big)\Big(1-\frac{\sqrt{2}-1}{\sqrt{2}}a_k^*a_k\Big)a_k$$
$$+a_k^*\Big(1-\frac{\sqrt{2}-1}{\sqrt{2}}a_k^*a_k\Big)\Big(1-\frac{\sqrt{2}-1}{\sqrt{2}}a_j^*a_j\Big)a_j\Big\}+v_0\sum_{j=1}^{L}a_j^{*2}a_j^2\,,\qquad v_0=\infty. \tag{2·15}$$

For $S=3/2$

$$\mathcal{H}=-\frac{9}{4}LzJ'-\frac{3}{2}LH+(3zJ'+H)\sum_{j=1}^{L}a_j^*a_j-2J'\sum_{<jk>}a_j^*a_ja_k^*a_k$$
$$-3J\sum_{<jk>}\Big\{a_j^*\Big(1-\frac{\sqrt{3}-\sqrt{2}}{\sqrt{3}}a_j^*a_j+\frac{\sqrt{3}-2\sqrt{2}+1}{2\sqrt{3}}a_j^{*2}a_j^2\Big)$$
$$\times\Big(1-\frac{\sqrt{3}-\sqrt{2}}{\sqrt{3}}a_j^*a_k+\frac{\sqrt{3}-2\sqrt{2}+1}{2\sqrt{2}}a_k^{*2}a_k^2\Big)a_k+\text{c. c.}\Big\}$$
$$+v_0\sum_{j=1}^{L}a_j^{*3}a_j^3\,,\quad v_0=\infty. \tag{2·16}$$

Van Kranendonk[9,10] has derived (2·14), rewritten it in the coordinate representation of N Bose particles and calculated the thermodynamic properties applying the usual Virial development of the theory of low density gases.

§ 3. Antiferromagnetism

The case to be considered in this section is the case where the Hamiltonian describing the system is given by (2.1) with negative J and J', the lattice of the system is constituted of two sublattices A and B and any lattice site of A is surrounded by sites of B and vice versa. Then our formulation of the previous section gives us as the equivalent Hamiltonian (2·13). Hence, for instance, for $S=1$ 2 and $H=0$ this case corresponds to a Bose lattice gas with negative mass and repulsive potential energy and the lowest eigenvalue E_0 of the Hamiltonian (2·13) is expected to be of the form a or b, of Fig. 1 as a function of N, L: It is a decreasing function of N L when N L is

[*] Mannari[14] has quite recently published a paper in which he assumes that the Heisenberg model of a ferromagnetic can be represented by a non-ideal Fermi lattice gas. The Hamiltonian assumed by him is equal to the one obtained by taking the Bose operators, a_j and a_i^*, in the Hamiltonian (2·14) to be the Fermi operators; which is justifiable for one dimensional case, because of the presence of a hard core in the interaction (cf. e.g. T. Morita: *Busseiron-Kenkyu* No. 89 (Oct. 1955), p. 12).

small and symmetric around $N/L=1/2$.* Hence the ground state is expected to be either the state with $N=L/2$, case *a*, or at least a state with $N\sim L/2$, case *b*. Therefore, if we are to treat the antiferromagnetic case by using the Hamiltonian (2·13), the problem becomes that of high density fluid even for very low temperatures, and no powerful method to treat the latter is known. The merit of (2·13) compared with the Hamiltonian which will be derived below rests on the fact that the former commutes with $\sum_{j=1}^{L} a_j^* a_j$.

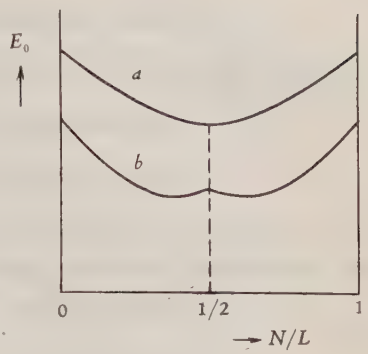


Fig. 1

Now, we turn to the Hamiltonian which is believed to be useful for the treatment of antiferromagnetics in very low temperatures. As we started in the previous section with the Hamiltonian due to Holstein and Primakoff, we will now start with the Hamiltonian due to Kubo.

We are to write the Hamiltonian in terms of the Bose operators a_j^*, a_j related to the lattice sites of A and b_k^*, b_k related to the lattice sites of B. We notice here the correspondences

$$\begin{aligned} S_j^+ &\sim (2S)^{1/2}\left(1-\frac{a_j^* a_j}{2S}\right)^{1/2} a_j \sim (2S)^{1/2}\sum_{n=0}^{2S-1}\frac{c_n}{n!} a_j^{*n} a_j^n a_j \\ S_j^- &\sim (2S)^{1/2} a_j^*\left(1-\frac{a_j^* a_j}{2S}\right)^{1/2} \sim (2S)^{1/2} a_j^* \sum_{n=0}^{2S-1}\frac{c_n}{n!} a_j^{*n} a_j^n \\ S_{jz} &\sim S - a_j^* a_j \end{aligned} \qquad (3\cdot 1)$$

related to a lattice site j of A, and

$$\begin{aligned} S_k^- &\sim (2S)^{1/2} b_k^*\left(1-\frac{b_k^* b_k}{2S}\right)^{1/2} \sim (2S)^{1/2} b_k^* \sum_{n=0}^{2S-1}\frac{c_n}{n!} b_k^{*n} b_k^n \\ S_k^- &\sim (2S)^{1/2}\left(1-\frac{b_k^* b_k}{2S}\right)^{1/2} b_k \sim (2S)^{1/2}\sum_{n=0}^{2S-1}\frac{c_n}{n!} b_k^{*n} b_k^n b_k \\ S_{kz} &\sim -S + b_k^* b_k \end{aligned} \qquad (3\cdot 2)$$

related to a lattice site k of B: Here the correspondence $\sim$ is in the same sense as those of (2·5) to S_j^+, S_j^- and S_{jz} in the previous section. Our eigenvalue problem becomes to be that of the Hamiltonian (2·1) with (3·1) and (3·2) in the restricted Hilbert space

---

* If there were no interactions, a particle with negative mass would have less energy when it has larger momentum, and then all particles in system would have the largest possible momentum.

Further, the lowest energy for a given number of particles is symmetric around $N/L=1/2$ because it should not depend on the choice of the positive direction of z axis.

$$\Pi_{j=1}^{L/2}\ (a_j^*)^{n_j}\Pi_{k=1}^{L/2}(b_k^*)^{n_k}\Psi_0 \qquad (n_j,\ n_k=0,\ 1,\ \cdots,\ 2S), \tag{3·3}$$

where the basic state Ψ_0 satisfying

$a_j\Psi_0=0$ for all j on the sublattice A and

$b_k\Psi_0=0$ for all k on the sublattice B,

is the state where all the spins on the sublattice A direct to the positive z direction and on the sublattice B to the negative z direction.

As in the previous section, we transform the problem to that in the extended Hilbert space

$$\Pi_{j=1}^{L/2}(a_j^*)^{n_j}\Pi_{k=1}^{L/2}(b_k^*)^{n_k}\Psi_0 \qquad (n_j,\ n_k=0,\ 1,\ \cdots,\ \infty) \tag{3·4}$$

by adding to our Hamiltonian a term

$$\mathcal{H}'=v_0\sum_{j=1}^{L/2}a_j^{*2S+1}a_j^{2S+1}+v_0\sum_{k=1}^{L/2}b_k^{*2S+1}b_k^{2S+1}, \qquad v_0=\infty, \tag{3·5}$$

which is infinite outside the original Hilbert space and zero inside it. That is, we have the Hamiltonian

$$\begin{aligned}\mathcal{H}=&-Lz\mathcal{S}^2|J'|+(2z\mathcal{S}|J'|+H)\sum_{j=1}^{L/2}a_j^*a_j+(2z\mathcal{S}|J'|-H)\sum_{k=1}^{L/2}b_k^*b_k-2|J'|\sum_{\langle jk\rangle}a_j^*a_jb_k^*b_k\\&+2\mathcal{S}|J|\sum_{\langle jk\rangle}'\Big\{a_j^*\sum_{n=0}^{2S-1}\frac{c_n}{n!}a_j^{*n}a_j^nb_k^*\sum_{m=0}^{2S-1}\frac{c_m}{m!}b_k^{*m}b_k^m+\sum_{n=0}^{2S-1}\frac{c_n}{n!}a_j^{*n}a_j^n\sum_{m=0}^{2S-1}\frac{c_m}{m!}b_k^{*m}b_k^m\Big\}\\&+v_0\sum_{j=1}^{L/2}a_j^{*2S+1}a_j^{2S+1}+v_0\sum_{k=1}^{L/2}b_k^{*2S+1}b_k^{2S+1}, \qquad v_0=\infty,\end{aligned} \tag{3·6}$$

in the Hilbert space (3·4) as an equivalent to the original Hamiltonian (2·1) in the original Hilbert space (2·7′).

For $\mathcal{S}=1/2$, it gives

$$\begin{aligned}\mathcal{H}=&-\frac{1}{4}Lz|J'|+(z|J'|+H)\sum_j'a_j^*a_j+(z|J'|-H)\sum_k'b_k^*b_k-2|J'|\sum_{\langle jk\rangle}'a_j^*a_jb_k^*b_k\\&+|J|\sum_{\langle jk\rangle}'(a_j^*b_k^*+a_jb_k)+v_0\sum_j'a_j^*a_j^*a_ja_j+v_0\sum_k'b_k^*b_k^*b_kb_k, \qquad v_0=\infty.\end{aligned} \tag{3·7}$$

This Hamiltonian does not commute with $\sum_{j=1}^{L/2}a_j^*a_j$, $\sum_{k=1}^{L/2}b_k^*b_k$ or their sum. As the result, the situation is not so simple as is the case for the ferromagnetics where the simple interpretation as the Bose lattice gas is possible.

§ 4. Ferrimagnetism

In this section, we treat a system of L lattice sites, which consist of two sublattices,

A and B, and on which the sites of A are occupied by a spin S_A and of B by a spin S_B and the site A is surrounded by sites of B and vice versa. The cases of positive J and J' and negative J and J' are treated seperately in 4*a*. and 4*b*. Only the results will be presented, as the derivations will be obvious from the preceding sections.

4a. Case of positive J and J'.

$$\mathcal{H} = -L z S_A S_B J' - \frac{1}{2} LH(\mu_A S_A + \mu_B S_B)$$

$$+ (2zS_B J' + \mu_A H) \sum_{j=1}^{L/2} a_j^* a_j + (2zS_A J' + \mu_B H) \sum_{k=1}^{L/2} a_k^* a_k - 2J' \sum_{<jk>} a_j^* a_j a_k^* a_k$$

$$- 2\sqrt{S_A S_B}\, J \sum_{<jk>} \Bigg[a_j^* \sum_{n=0}^{2S_A - 1} \frac{c_n}{n!} a_j^{*n} a_j^n \sum_{m=0}^{2S_B - 1} \frac{c_m}{m!} a_k^{*m} a_k^m a_k + \text{c. c.} \Bigg]$$

$$+ v_0 \sum_{j=1}^{L/2} a_j^{*2S_A+1} a_j^{2S_A+1} + v_0 \sum_{k=1}^{L/2} a_k^{*2S_B+1} a_k^{2S_B+1}, \qquad v_0 = \infty. \tag{4·1}*$$

4b. Case of negative J and J'.

$$\mathcal{H} = -L z S_A S_B |J'| - \frac{1}{2} LH(\mu_A S_A - \mu_B S_B)$$

$$+ (2zS|J'| + \mu AH) \sum_{j=1}^{L/2} a_j^* a_j + (2zS_A |J'| - \mu_B H) \sum_{k=1}^{L/2} b_k^* b_k - 2|J'| \sum_{<jk>} a_j^* a_j b_k^* b_k$$

$$+ 2\sqrt{S_A S_B}\, |J| \sum_{<jk>} \Bigg[a_j^* \sum_{n=0}^{2S_A - 1} \frac{c_n}{n!} a_j^{*n} a_j^n b_k^* \sum_{m=0}^{2S_B - 1} \frac{c_m}{m!} b_k^{*m} b_k^m + \text{c. c.} \Bigg]$$

$$+ v_0 \sum_{j=1}^{L/2} a_j^{*2S_A+1} a_j^{2S_A+1} + v_0 \sum_{k=1}^{L/2} b_k^{*2S_B+1} b_k^{2S_B+1}, \qquad v_0 = \infty. \tag{4·2}$$

§ 5. Partition function for a Hamiltonian with infinite potential

We have introduced an infinite potential to the Hamiltonian to restrict the Hilbert space in the correct region. When we calculate the partition function for the Hamiltonian, containing an infinite term which restricts the Hilbert space, we may make use of the following technique.

The procedure used in §§ 2, 3 and 4 is a special application of the theorem that the eigenvalue problem of $\mathcal{H}_0$ in a Hilbert space R_0 is equivalent to the eigenvalue problem of $\mathcal{H}_0 + \mathcal{H}'$ in a Hilbert space $R = R_0 + R'$ when $\mathcal{H}'$ is infinity in R' and is zero in R_0: i.e.

* μ_A and μ_B denote g-factors of electrons on the respective lattice site, multiplied by the Bohr magneton. For the cases of ferro- and antiferro-magnetism, we have had this factor included in H.

$$\mathcal{H}'=v_0\sum_{\mu}|\psi_\mu><\psi_\mu|, \qquad v_0\to\infty, \tag{5·1}$$

where ψ_μ is a complete set of the Hilbert space R'. As the consequence, the partition functions of $\mathcal{H}_0$ in space R_0 and of $\mathcal{H}$ in space R coincide. When we write complete setes of R_0 and R as Ψ_μ, Φ_μ, we have two formulae for the partition function Z

$$Z=\mathrm{T}_{\mathrm{r}}^{(R)}[e^{-\beta H}]=\sum_{\mu}(\Phi_\mu|e^{-\beta H}|\Phi_\mu), \tag{5·2}$$

$$Z=\mathrm{T}_{\mathrm{r}}^{(R_0)}|e^{-\beta H_0}|=\sum_{\mu}(\Psi_\mu, e^{-\beta H_0}\Psi_\mu)$$

$$=\sum_{\mu,\lambda}(\Phi_\mu|e^{-\beta H_0}|\Phi_\lambda)(\Phi_\lambda|E_{R_0}|\Phi_\mu), \tag{5·3}$$

$$E_{R_0}=\sum|\Psi_\mu)(\Psi_\mu|. \tag{5·4}$$

We may use any of eq. (5·2) and eq. (5·3) for the calculation of Z. In the calculation by use of (5·3), the presence of the projection operator may introduce some difficulties. However, when we calculate the partition function by using some methods, in the process of which intermediate states appear (as the method of path integral), the intermediate states are taken over R without restriction, while they are restricted inside R_0 when (5·2) is used.

It is noted that, when $\mathcal{H}$ is given, $\mathcal{H}_0$ is ambiguous by a term of the form

$$\mathcal{H}''=\sum_{\mu}v_\mu|\Psi_\mu)(\Psi_\mu|, \qquad v_\mu>-\infty \tag{5·5}$$

which does not change the eigenvalue problem because

$$\mathcal{H}''\Psi_\mu=0 \qquad \text{for all } \mu. \tag{5·6}$$

Hence we may add to $\mathcal{H}_0$ in (5·3) a term $\mathcal{H}''$ of the form (5·5).

As an example, we consider an isotropic ferromagnetic with spin 1/2. The Hamiltonian is $\mathcal{H}$ given by (2·14) with $J=J'$. The most natural choice of $\mathcal{H}_0$ will be the one, $\mathcal{H}$ minus $v_0\sum_{j=1}^{N}a_j^*a_j^*a_ja_j$. From the result, we find that Dyson has calculated (5·3) with $\mathcal{H}_0$

$$\mathcal{H}_0=\mathcal{H}-v_0\sum_{j=1}^{N}a_j^*a_j^*a_ja_j+J\sum_{<jk>}a_j^*a_k^*(a_j^2+a_k^2)\ ; \tag{5·7}$$

the last term has the property (5·6).

On the other hand, Van Kranendonk has used eq. (5·2) directly. The discrepancies of their calculations are attributed to their ways of approximation in the calculations. This point will be investigated in the following paper.

For the case of $S\neq 1/2$, the discrepancy of eq. (57) of Dyson's paper and our eq. (2·13) minus $v_0\sum_{j=1}^{N}a_j^{*2S+1}a_j^{2S+1}$ will deserve further investigations. Our Hamiltonian is Hermitian while eq. (5·7) of Dyson's is not so. But in our case, the presence of many body interactions makes the problem more difficult.

§ 6. Conclusion

We have obtained the interacting Bose lattice gas equivalent to the Heisenberg model of ferro-, antiferro- and ferri-magnetism. When we neglect the interactions, the Hamiltonian obtained reduces to that of spin wave. With our Hamiltonian obtained we can apply the theories of low density gases to the problems of magnetism which has been applied to the ferromagnetism with spin 1/2 by Van Kranendonk and Dyson. The interaction is two-body force when $S=1/2$, three-body force when $S=1$ and generally $(4S-1)$ body force for $S\geq 1$. When we are interested in a state where the magnetization is far from the saturations, the presence of infinite potential makes the method unapplicable at present, because the problem reduces to that of high density fluid for which we know no good method at present.

In the last section, we have presented a technique to calculate the partition function for Hamiltonian with an infinite potential. In connection with this, further investigation about the choice of $\mathcal{H}_0$ is desired.

Although there may remain several points to be investigated in connection with our Hamiltonians, we have obtained the Hamiltonians which are equivalent to those of the anisotropic spin systems and which have the form of finite power series of the Bose operators and for which we have no such problem as the inevitable cut-off of infinite power series reults in the spreading out of the Hilbert space as was the case with the Hamiltonian of Holstein and Primakoff for ferromagnetics and the Hamiltonian of Kubo for antiferromagnetics.

The author wishes to express his sincere thanks to Dr. K. Hiroike for his helpful discussions.

Appendix: Calculation of c_n's

c_n's are defined by eq. (2·9):

$$\left(1-\frac{a^*a}{2S}\right)^{1/2}=\sum_{n=0}^{\infty}\frac{c_n}{n!}a^{*n}a^n\,. \tag{A·1}$$

Operating this to Φ_0, $a^*\Phi_0$, $a^*a^*\Phi_0$, $\cdots$, we have

$$\Phi_0=c_0\Phi_0 \tag{A·2^0}$$

$$\left(1-\frac{1}{2S}\right)^{1/2}a^*\Phi_0=(c_0+c_1)\,a^*\Phi_0 \tag{A·2^1}$$

$$\left(1-\frac{2}{2S}\right)^{1/2}a^*a^*\Phi_0=\left(c_0+2c_1+\frac{2\cdot 1}{2!}c_2\right)a^*a^*\Phi_0 \tag{A·2^2}$$

$\cdots\cdots\cdots\cdots$

Generally, operating it to $a^{*n}\Phi_0$, we have

$$\left(1-\frac{n}{2S}\right)^{1/2}a^{*n}\Phi_0=\sum_{m=0}^{n}\binom{n}{m}c_m a^{*n}\Phi_0\,. \tag{A·2^n}$$

Solving (A·2) successively, we obtain*

$$\begin{aligned} c_0 &= 1, \\ c_1 &= -1 + (1 - 1/2S)^{1/2}, \\ c_2 &= 1 - 2(1 - 1/2S)^{1/2} + (1 - 1/S)^{1/2}, \\ &\cdots\cdots\cdots \end{aligned} \tag{A·3}$$

For $S=1/2$,

$$c_0 = 1, \quad c_1 = -1, \cdots\cdots$$

For $S=1$,

$$c_0 = 1, \quad c_1 = -1 + 1/\sqrt{2}, \quad c_2 = 1 - \sqrt{2}, \cdots\cdots$$

For $S=3/2$,

$$c_0 = 1, \quad c_1 = -1 + \sqrt{2/3}, \quad c_2 = 1 - 2\sqrt{2/3} + 1/\sqrt{3}, \cdots\cdots$$

References

1) J. Van Kranendonk and J. H. Van Vleck; Rev. Mod. Phys. **30** (1958), 1.
2) F. Bloch, Z. Physik **61** (1930), 206; **74** (1932), 295.
3) T. Holstein and H. Primakoff, Phys. Rev. **58** (1940), 1098.
4) R. Kubo, Phys. Rev. **87** (1952), 568.
   T. Nagamiya, K. Yoshida and R. Kubo, Adv. Phys. **4** (1955), 1.
5) P. W. Anderson, Phys. Rev. **86** (1952), 694.
6) T. Morita, Prog. Theor. Phys. **18** (1957), 462; See also Ref. 7.
7) T. Matsubara and H. Matsuda, Prog. Theor. Phys. **16** (1956), 569.
8) H. Matsuda and T. Matsubara, Prog Theor. Phys. **17** (1957), 19.
   H. Matsuda, Prog. Theor. Phys. **18** (1957), 357.
9), 10) J. Van Kranendonk, Physica **21** (1955), 749 and 925.
11), 12) F. J. Dyson, Phys, Rev. **102** (1956), 1217 and 1230.
13) Footnote 13 of Ref. 3.
14) I. Mannari, Prog. Theor. Phys. **19** (1958), 201.

* The fact that c_n's for $n>2S$ are complex will do no harm because what used in the context is only that both sides of (A·1) have the same effect when operated to $a^{*n}\Phi_0$ with $n\leq 2S$.

Progress of Theoretical Physics, Vol. 20, No. 5, November 1958

Nitrogen Induced Nuclear Reactions in Sodium

Ryuzo NAKASIMA

Research Institute for Atomic Energy, Osaka City University, Osaka

(Received June 20, 1958)

The nuclear reactions $^{23}Na(^{14}N;\alpha p)^{32}P$ and ^{23}Na $(^{14}N;\alpha 2p)$ ^{31}Si are analysed on the basis of the compound nucleus model. The ratio of the cross sections in both reactions at 26 Mev bombarding ion energy seems to be fairly well reproduced in terms of the usual compound nucleus picture, but the absolute values of these cross sections are anomalously small. Taking account of the effects of depressed Coulomb barrier height, we obtain the right order of magnitudes for the excitation cross sections of both reactions. According to the present investigation, it seems to be rather premature to argue definitely for the presence of buckshot type nuclear reaction, but suggestions and discussions concerning the reaction mechanism of the heavy ion induced nuclear reactions are given.

§ 1. Introduction

A number of experiments on heavy ion induced nuclear reactions so far presented are classified into two categories according to the type of the reaction mechanism. One is the multiple evaporation type nuclear reaction proceeding through the compound nucleus formation process, and it leads to the residual nuclei a few mass units lighter than the compound nucleus. The other is the nucleon transfer reaction which results in reaction products one mass unit lighter or heavier than the initial nuclei involved in the reaction. The latter process has been treated theoretically by Breit and Ebel[1] and it has been clarified that the reaction mechanism is closely related to nuclear structure. On the other hand, we have analysed in a previous paper[2] the multiple evaporation process in $^{14}N+^{12}C$ reaction in which it has been pointed out that the gross feature of the excitation cross sections is fairly well interpreted in terms of the compound nucleus picture. Although the compound nucleus formalism may be applied to the interpretation on gross feature of the cross sections of heavy ion induced nuclear reactions, some kinds of direct interaction must not be neglected to obtain better agreement with experiment. In the previous work[2], it has been shown that the partial capture process may contribute in the case of $^{12}C(^{14}N\,;2\alpha)^{18}F$ reaction. This kind of direct interaction may be interpreted as the so-called buckshot type reaction mechanism since the nuclei involved in the reaction are broken-up already before they come into contact.

The possibility of buckshot type nuclear reaction is discussed in some detail by Chackett, Fremlin and Walker[3] for $^{14}N+^{27}Al$ nuclear reaction. According to their suggestion, the higher the Coulomb barrier is, the more pronounced is the buckshot type process, and when the buckshot type process occurs the capture cross section is larger than the single bullet process. But it seems to be difficult to account for the effect only by the buckshot type nuclear reaction, because the effect of fall-down of the Coulomb barrier is not taken in

their discussion. In fact, in our preliminary report[4] in which the reactions ^{23}Na(^{14}N ; αp)^{32}P and ^{23}Na(^{14}N ; $\alpha 2p$)^{31}Si were analysed based on the compound nucleus model, we have pointed out that the ratio of the cross sections in both reactions at 26 Mev bombarding ion energy agrees with the experimental results[5], but the absolute values of these cross sections are largely scattered from those of experiments. This discrepancy seems to be explainable by the fall-down effect of the Coulomb barrier rather than by the buckshot type nuclear process.

The importance of the fall-down effect of the Coulomb barrier height has been pointed out already by Breit, Hull and Gluckstern[6] in the case of heavy ion bombardment on multi-charged target nucleus. They have considered the distortion effects in nuclei involved in the reaction on the basis of the liquid drop molel. Another investigation on this subject has been made by Kikuchi[7] who has shown in analysing the reaction ^{27}Al(n ; α)^{24}Na that the diffuseness of nuclear surface makes the penetration of the Coulomb potential barrier easier.

In the present work, the experimental results[5] of ^{23}Na(^{14}N ; αp) ^{32}P and ^{23}Na(^{14}N ; $\alpha 2p$) ^{31}Si reactions are analysed again by means of the compound nucleus model slightly modified by taking into account the depression of the potential barrier height. Our interest lies chiefly in the mechanism of the nuclear reaction induced by heavy ions, and if the effect of the buckshot type nuclear process or of the distortion of nuclei exists, it will be more appreciable in ^{14}N + ^{23}Na reactions than in ^{14}N − ^{12}C reactions analysed previously.

§ 2. Excitation cross sections

In the calculation based on the compound nucleus model, all decay processes energetically possible must be taken into account as the competitive processes. The treatment for the calculation of excitation cross section of multiple evaporation nuclear reaction is shown in a previous paper[8], in which three particles or gamma rays emissions are taken into account as the competitive processes, so that we will follow the procedure described there.

The cross section of the nuclear reaction (l ; i, j, k) is expressed, in general, as

$$\sigma(l; i, j, k) = \sigma_C^{(l)}(E_l) \cdot K_{ijk} / \sum_{m,n,p} K_{mnp} = \sigma_C^{(l)}(E_l) \cdot \mathfrak{P}_{ijk}, \qquad (1)$$

where $\sigma_C^{(l)}(E_l)$ is the cross section for the formation of a compound nucleus by the impinging particle l with incident energy E_l, and $\mathfrak{P}_{ijk}$ is the decay probability for having emission i, j and k from the compound nucleus. In the calculation of $\mathfrak{P}_{ijk}$, the constant a appearing in the level density formula $w(E) = c \cdot \exp(2\sqrt{aE})$ is chosen as 0.42 Mev^{-1} for ^{32}P according to the previously proposed relationship[9]. The effect of the depressed Coulomb barrier height must be taken into acccount in the evaluation of eq. (1), not only for $\sigma_C^{(l)}(E_l)$ but for $\mathfrak{P}_{ijk}$.

The distortion effects in nuclei, suggested by Breit et al[6], is taken into account by assuming the form of the nuclear deformation as

$$\delta r(\theta) = [\alpha_0 + \alpha_2 P_2(\cos\theta)],$$

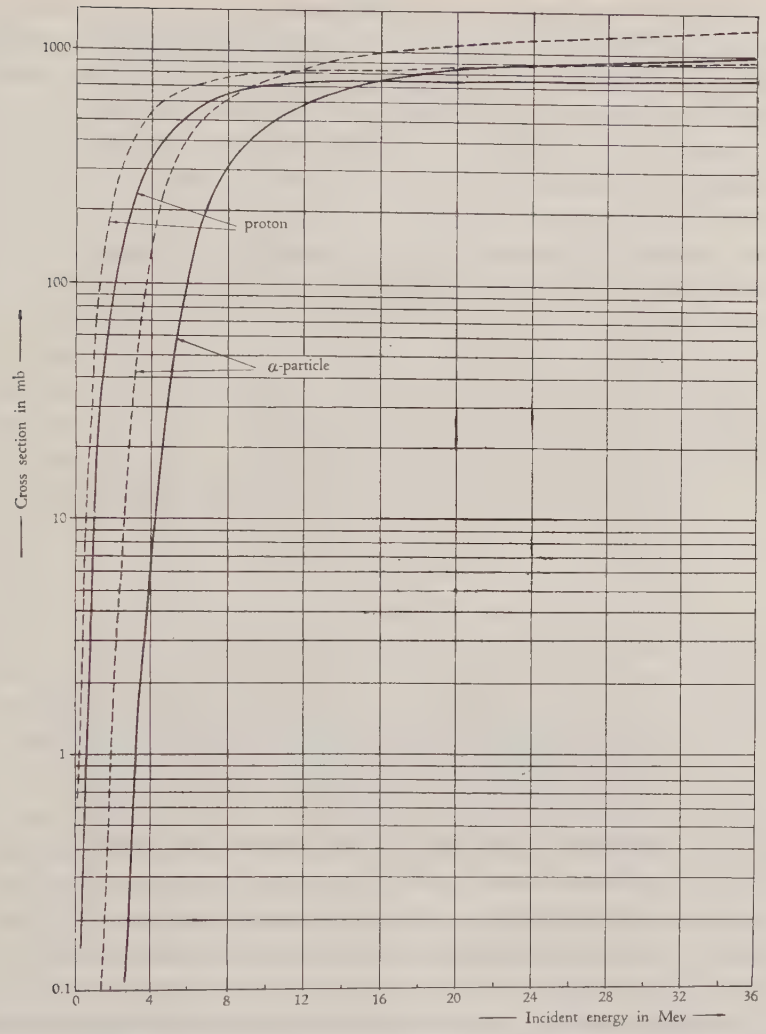


Fig. 1. A typical illustration of the compound nucleus formation cross sections by proton and alpha-particle on a phosphorus target (Z=15). The solid curves refer to the calculated cross sections using the usual Coulomb barrier, and the dashed curves are those using the depressed Coulomb barrier.

where α_2 must be determined from the condition of potential minimum after the Coulomb energy is averaged over θ and r. Taking account of their consideration into the present case, one may find that the effective Coulomb barrier height is about 90% of the usual one, and for the purpose of the present investigation it is not quite effective. On the other hand, the effect of the diffused boundary of the nuclei results in the ratio of the effective Coulomb barrier height to the usual one of about 2/3. In this estimate we supposed the potential as follows:

$$V(r) = V_n(r) + V_c(r) \qquad \text{for } r > R,$$

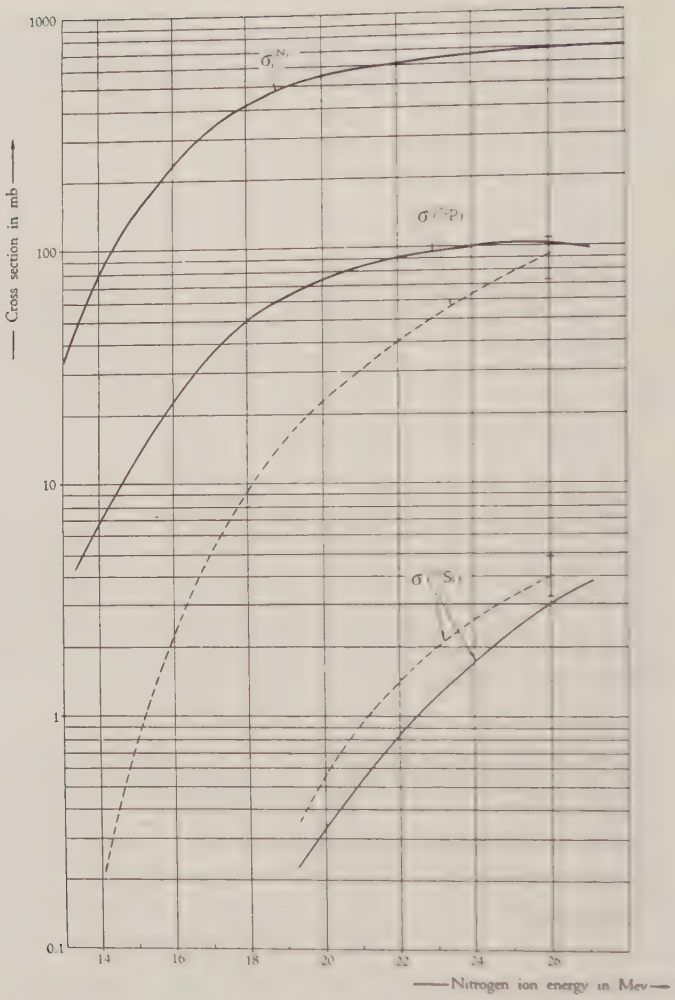


Fig. 2. Comparison between the calculated and experimental results for the nitrogen induced nuclear reactions on sodium. The solid curves refer to the calculated cross sections and the dashed curves are those of the experiments. The compound nucleus formation cross section for $^{14}N+^{23}Na$ is shown also.

where V_n and V_c are the nuclear potential and the Coulomb potential, respectively, which are given by

$$V_n(r) = -V_0 \cdot \left[1 + \exp\left(\frac{r - R}{a} \right) \right]^{-1}$$

$$V_c(r) = ZZ'e^2 / r.$$

Then one may find

$$B_{crit} \; B_c \sim 1 - 1.7 \; R,$$

where R is the nuclear radius in unit of 10^{-13}cm, and expressed as $1.5 \times A^{1/3}$. With the

effective Coulomb barrier height, diminished by the above mentioned effects, the cross sections of compound nucleus formation are evaluated for proton, alpha-particle and nitrogen ion, and these cross sections are illustrated in Fig. 1 for proton and alpha-particle. Furthermore, for the case of neutron, the compound nucleus formation cross section is affected also by the rounded-off nuclear potential as was pointed out by Kikuchi[7].

We make the calculation of the excitation functions based on the compound nucleus picture with above mentioned modification. In Fig. 2 the experimental and the calculated excitation cross sections under consideration are illustrated. Also the cross section of compound nucleus formation for impinging nitrogen ion on sodium target is illustrated in this figure. Though the cross section at 26 Mev using the conventional theory is about 430 mb, the present calculation with the diminished Coulomb barrier gives about 700mb for this cross section. In the evaluation of the compound nucleus formation cross section when incident nitrogen ion collides with sodium target, the channel radius is taken somewhat smaller than the sum of the radii of ^{14}N and ^{23}Na nuclei and larger than the radius of the compound nucleus ^{37}A.

§ 3. Discussions

From Fig. 2 it is seen that the overall behavior of the excitation cross sections of both reactions is fairly well reproduced by the present calculations. It must be noted, however, that the discrepancies in these curves seem to suggest us many imports.

In the first place, if the calculated curve of $\sigma(^{32}\mathrm{P})$ is shifted to right by about 4 Mev in scale of incident energy it coincides with that of experiment, and for the case of $\sigma(^{31}\mathrm{Si})$ it has to be shifted to left by about 1.5 Mev. Therefore, provided that the separation energies used here are unsuitable, a possibility of improvement is expected. In the present calculation, however, the separation energies were calculated from the mass excess data listed by Bethe and Morrison[9], and thus it is unlikely to give rise to a mistake of about 5.5 Mev in difference of separation energies for both cases. Therefore, we are obliged to consider a special mechanism for nuclear reactions induced by heavy ion bombardments in order to explain these discrepancies. As a different point of view, if we compare the calculated excitation function of $^{23}\mathrm{Na}(^{14}\mathrm{N}\,;\,\alpha p)^{32}\mathrm{P}$ with that[2] of $^{12}\mathrm{C}(^{14}\mathrm{N}\,;\,\alpha)^{22}\mathrm{Na}$, a close similarity in the discrepancies may be found. In the previous case, the discrepancy is discussed in terms of the isotopic spin selection rule in nuclear reactions but in the present case it is premature to make such interpretation because of incompletion of our knowledge on the isotopic spin assignment to the complex nuclei taking part in the reaction.

Next, let us consider the possibility of occurrence of the buckshot type nuclear process. Recent experiment[10] on the angular distribution of reaction products of nitrogen induced nuclear reaction seems to be a typical example indicating the existence of buckshot type nuclear process though the original author makes no mention of buckshot type reaction. Zucker has measured the angular distributions of protons and alpha-particles from the reactions $^{27}\mathrm{Al}(^{14}\mathrm{N}:p)^{40}\mathrm{K}$ and $^{27}\mathrm{Al}(^{14}\mathrm{N}\,;\,\alpha)^{37}\mathrm{A}$. The protons from the former reaction

are emitted isotropically whereas the angular distribution of alpha-particles shows the definite forward peak. Since the type of this experimental result is not multiple evaporation type, one cannot emphasize immediately the existence of the buckshot type reaction in the multiple evaporation type nuclear reactions. However, taking account of the fact that the emission of alpha-particles occur frequently in most cases of the multiple evaporation type reactions, we may expect the break-up process of impinging nitrogen ion such as $^{14}N \rightarrow \alpha + {}^{10}B$. Here we assume that the alpha-particle is scattered directly, while the boron nucleus is captured in the target nucleus of sodium to form a compound nucleus ^{33}S, and one proton and gamma ray or two protons are evaporated from the compound nucleus ^{33}S successively. Estimating the branching ratio $\sigma(^{32}P)/\sigma(^{31}Si)$ in this case, one finds smaller values compared with the calculated results for the complete compound nucleus formation process (Section 2) over the range of the incident energy under consideration. This tendency is just what we look for, but unfortunately we cannot make quantitative discussions at the present time since it is not known what amount of the incident energy is shared to the compound nucleus produced by the partial capture of boron nucleus. Therefore, it is desired to design the experiments so as to measure the angular distribution and energy spectrum of alpha-particles from the multiple evaporation type nuclear reactions as in the recent experiment[10] on heavy ion induced nuclear reaction with one kind of reaction product.

Finally, the compound nucleus formation cross section for impinging nitrogen ion on a sodium target is quite similar to calculated excitation cross section for $^{23}Na(^{14}N : \alpha p)^{32}P$ reaction, as is seen from Fig. 2. It stands for the constancy of $\mathfrak{P}_{s,\gamma}$ in eq. (1) over this large energy region. Similar situation was pointed out by Reynolds and Zucker[11] for the case of $^{9}Be(^{14}N ; \alpha n)^{18}F$ reaction in which the similarity lies between the compound nucleus formation cross section and the experimental excitation function. Accordingly, it does not seem to be so surprising a fact that, in the case of $^{9}Be(^{14}N ; \alpha n)^{18}F$ reaction, the ratio of the reaction cross section to the total cross section (cross section for compound nucleus formation) remains nearly constant over the large incident energy region.

The author would like to express his gratitude to Professor K. Husimi for his guidance and encouragement. He is also indebted to Messrs. T. Kammuri and S. Yamasaki for many helpful discussions.

References

1) G. Breit, Phys. Rev. **102** (1956), 549.
G. Breit and M. E. Ebel, Phys. Rev. **103** (1956), 679.
G. Breit and M. E. Ebel, Phys. Rev. **104** (1956), 1030.
2) R. Nakasima, Y. Tanaka and K. Kikuchi, Prog. Theor. Phys. **15** (1956), 574.
3) K. F. Chackett, J. H. Fremlin and D. Walker, Phil. Mag. **45** (1954), 173.
4) R. Nakasima, Prog. Theor. Phys. **19** (1958), 743.
5) M. L. Halbert, T. H. Handley and A. Zucker, Phys. Rev. **104** (1956), 115.
6) G. Breit, M. H. Hull, Jr. and R. L. Gluckstern, Phys. Rev. **87** (1952), 74.
7) K. Kikuchi, Prog. Theor. Phys. **17** (1957), 643.
8) R. Nakasima and K. Kikuchi, Prog. Theor. Phys. **14** (1955), 126.
9) H. A. Bethe and P. Morrison, "*Elementary Nucleay Theory*", John Wiley & Sons (1956).
10) A. Zucker, Nuclear Physics **6** (1958), 420.
11) H. L. Reynolds and A. Zucker, Phys. Rev. **100** (1955), 226.

Progress of Theoretical Physics, Vol. 20, No. 5, November 1958

The Interaction between Two Normal Helium Atoms

—— *A Consistent Treatment* ——

Tasuke HASHINO and Sigeru HUZINAGA

General Education Department, Kyushu University, Otsubo-machi, Fukuoka

(Received July 14, 1958)

We report here a consistent variation calculation of the interaction energy between two normal helium atoms from the repulsive region to the van der Waals region. The LCAO MO method within single configuration approximation is used and two variation parameters are included in the wave function in such a way that the bonding and the anti-bonding orbitals can be deformed separately. The result of the calculation is satisfactory.

Introduction

The calculation of the interaction energy between two normal helium atoms is one of the most fundamental problems in the quantum theory of molecular systems. It has received various kinds of treatment. If we try to tackle this problem by the method of the molecular orbital in the single configuration approximation, it is natural to adopt one bonding orbital $\varphi_g(\sigma_g)$ and one antibonding orbital $\varphi_u(\sigma_u)$ and assign two electrons with opposite spins to each of these orbitals, according to the Pauli principle. Thus, the wave function for the total system which consists of two normal helium atoms, separated by finite distance, may be written as

$$\Psi = (1/24)^{1/2} \sum_{\lambda} (-1)^{\lambda} P_{\lambda} [\varphi_g(1)\alpha(1)\varphi_g(2)\beta(2)\varphi_u(3)\alpha(3)\varphi_u(4)\beta(4)], \qquad (1)$$

where P_λ is the operation permuting the electron coordinates, 1, 2, 3, 4, and α and β are the usual spin eigenfunctions. The functions, φ_g, φ_u, α and β, are all supposed to be normalized to unity.

Griffing and Wehner[1] adopted the so-called LCAO approximation for φ_g and φ_u;

$$\varphi_g = N_g(s_a + s_b) \quad ; \quad \sigma_g, \qquad (2)$$

$$\varphi_u = N_u(s_a - s_b) \quad ; \quad \sigma_u, \qquad (3)$$

where N_g and N_u are the normalization constants and

$$s_a = (Z^3/\pi)^{1/2} \exp(-Zr_a), \quad s_b = (Z^3/\pi)^{1/2} \exp(-Zr_b). \qquad (4)$$

Here Z is the parameter of the theory and r_a and r_b are the distances of the electron from nucleus a and b respectively. s_a and s_b are simple ls-functions, Z, the effective nuclear charge, is taken as $27/16$ in the calculation of Griffing and Wehner, according to a variational calculation on the atom. They made an attempt to improve the calcul-

ation by minimizing the energy with respect to the parameter Z for each internuclear distance, only to find that departures from the value $Z=27$ 16 bring about almost always no signifficant improvement. Then they concluded that "the variation of effective nuclear charge is unimportant in the He_2 problem from the view-point of decreasing the energy."

As is well known, the wave function Ψ given in (1) consisting of the molecular orbitals $\varphi_g \sim s_a + s_b$ and $\varphi_u \sim s_a - s_b$ can be recast into the following form:

$$\Psi \sim \sum_{\lambda} (-1)^{\lambda} P_{\lambda} [s_a(1)\alpha(1)s_a(2)\beta(2)s_b(3)\alpha(3)s_b(4)\beta(4)]. \tag{5}$$

Hence, in this case, the molecular orbital theory and the valence bond theory are equivalent in the first stage of approximation. s_a and s_b are the ls atomic orbitals given in (4). If we express them in the prolate spheroidal coordinates,

$$\begin{aligned} s_a &\sim \exp[-(ZR/2)\xi + (ZR/2)\eta], \\ s_b &\sim \exp[-(ZR/2)\xi - (ZR/2)\eta], \end{aligned} \tag{6}$$

where

$$\xi = (r_a + r_b)/R, \quad \eta = (r_a - r_b)/R,$$

or conversely

$$r_a = (R/2)(\xi + \eta), \quad r_b = (R/2)(\xi - \eta).$$

Sakamoto and Ishiguro[2] attempted a refinement in the valence bond scheme by adopting the modified atomic orbitals introduced by Inui[3] and Nordsieck[4], namely

$$\begin{aligned} s_a &\sim \exp(-\alpha\xi + \beta\eta), \\ s_b &\sim \exp(-\alpha\xi - \beta\eta), \end{aligned} \tag{7}$$

where α and β are the independent parameters. Note the relation between (7) and (6). In this way, they attempted to take into account the deformation of the atomic orbitals due to the approach of another atom. The result was not encouraging. The repulsive interaction energy of the He-He system obtained by Sakamoto and Ishiguro is lower than that of Griffing and Wehner but the difference between these two calculations is rather small, namely within 1 eV over nearly the whole range of the internuclear distance discussed there.

Recently, another alternative refinement of the treatment of Griffing and Wehner was put forward by one of the present authors[5]. The idea is simple. We use the following molecular orbitals;

$$\varphi_g \sim s_a + s_b \ ; \ \sigma_g, \tag{8}$$

$$\varphi_u \sim s_a' - s_b' \ ; \ \sigma_u, \tag{9}$$

where

$$\begin{aligned} s_a &= (Z^3/\pi)^{1/2} \exp(-Zr_a), \qquad s_b = (Z^3/\pi)^{1/2} \exp(-Zr_b), \\ s_a' &= (Z'^3/\pi)^{1/2} \exp(-Z'r_a), \qquad s_b' = (Z'^3/\pi)^{1/2} \exp(-Z'r_b). \end{aligned} \tag{10}$$

The point is that we take $Z \neq Z'$. This scheme is nothing but a slight generalization of that of Griffing and Wehner but the improvement brought about by this modification is found to be remarkable. The primary reason for this achievement is that the bonding orbital φ_g and the anti-bonding orbital φ_u can be deformed separately. The importance of such a freedom was first noted for the H_2^+ ion by Coulson[6] in 1937. Roughly speaking, the bonding orbital tends to contract inwardly and, on the contrary, the anti-bonding orbital wants to expand outwardly, these tendencies being understood most easily from the united-atom point of view. Various methods may be deviced to give such flexibilities to the molecular orbitals for the present problem.

The primary purpose of the LCAO MO calculation with $Z \neq Z'$[5] was to diminish the large discrepancy between theory and experiment at a small internuclear distance. This is the reason why we performed the calculation only at the distance of $1 \lesssim R \lesssim 2$ au. However, the force range which is physically most important and of direct interest may probably be the van der Waals region and the neighboring inside. We report in this paper a consistent quantum mechanical variation calculation of the interaction energy between two normal helium atoms from the repulsive region to the van der Waals attraction region.

Calculation and result

As a simple device to give necessary flexibilities to the bonding and the anti-bonding orbitals, we take φ_g and φ_u as follows:

$$\varphi_g = N_g[(s_a + s_b) + \lambda_g(p_a + p_b)]\,;\ \sigma_g, \tag{11}$$

$$\varphi_u = N_u[(s_a - s_b) + \lambda_u(p_a - p_b)]\,;\ \sigma_u, \tag{12}$$

where

$$\begin{aligned} s_a &= (1s\,\sigma)_a = (Z^3/\pi)^{1/2} \exp(-Zr_a), \\ s_b &= (1s\,\sigma)_b = (Z^3/\pi)^{1/2} \exp(-Zr_b), \\ p_a &= (2p\,\sigma)_a = (Z^5/\pi)^{1/2} r_a \cos\theta_a \exp(-Zr_a), \\ p_b &= (2p\,\sigma)_b = (Z^5/\pi)^{1/2} r_b \cos\theta_b \exp(-Zr_b). \end{aligned} \tag{13}$$

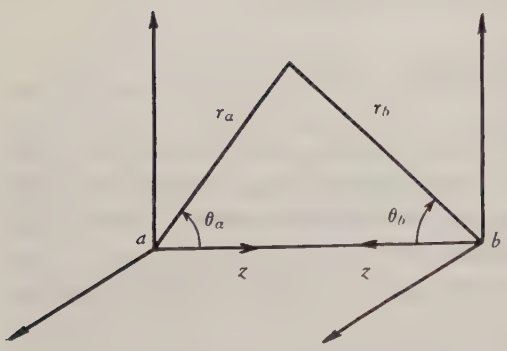


Fig. 1

Here, λ_g and λ_u are the variable parameters and N_g and N_u are the normalization factors which, needless to say, depend on λ_g and λ_u respectively. To simplify the evaluations of molecular integrals, a fixed value of $Z = 27/16$ is used commonly for $(1s\sigma)$ and $(2p\sigma)$ orbitals throughout the whole range of the internuclear separation.

The same kind of device for the orbital deformation was first adopted by Rosen[7] for H_2 in the valence bond theory and then

utilized by Dickinson[8] in molecular orbital treatment of H_2^-. Griffing and Wehner[1] have also made use of it in order to calculate the attraction energy between two normal helium atoms at one point near the minimum on the van der Waals energy curve. It is our aim to use the above orbitals to cover not only the van der Waals region but also the intermediate and the repulsive regions of the He-He interaction.

The wave function of the system is given in (1), namely,

$$\Psi = (1/24)^{1/2} \sum_{\lambda} (-1)^{\lambda} P_{\lambda} [\varphi_g(1)\alpha(1)\varphi_g(2)\beta(2)\varphi_u(3)\alpha(3)\varphi_u(4)\beta(4)].$$

Here φ_g and φ_u are given in (11) and (12). The energy operator for the system is (in atomic units*)

$$H = \sum_{i}^{4} H(i) + \sum_{i>j}^{4} (1/r_{ij}) + (4/R), \tag{14}$$

where

$$H(i) = -(1/2)\Delta_i - (2/r_{ai}) - (2/r_{bi}) .$$

Then the total energy of the system is

$$E(R) = (\Psi, H\Psi) = 2H_g + 2H_u + J_{gg} + J_{uu} + 4J_{gu} - 2K_{gu} + (4/R), \tag{15}$$

where

$$H_g = \int \varphi_g(1) H(1) \varphi_g(1) d\tau_1,$$

$$H_u = \int \varphi_u(1) H(1) \varphi_u(1) d\tau_1,$$

$$J_{gg} = \iint \varphi_g(1)\varphi_g(1)(1/r_{12})\varphi_g(2)\varphi_g(2) d\tau_1 d\tau_2,$$

$$J_{uu} = \iint \varphi_u(1)\varphi_u(1)(1/r_{12})\varphi_u(2)\varphi_u(2) d\tau_1 d\tau_2,$$

$$J_{gu} = \iint \varphi_g(1)\varphi_g(1)(1/r_{12})\varphi_u(2)\varphi_u(2) d\tau_1 d\tau_2,$$

$$K_{gu} = \iint \varphi_g(1)\varphi_u(1)(1/r_{12})\varphi_g(2)\varphi_u(2) d\tau_1 d\tau_2 .$$

As is easily seen, the expressions, $H_g N_g^2$, $H_u N_u^2$, $J_{gg} N_g^4$, $J_{uu} N_u^4$, $J_{gu} N_g^2 N_u^2$ and $K_{gu} N_g^2 N_u^2$, together with $1 N_g^2$ and $1 N_u^2$, are decomposed into polynomials of λ_g and λ_u. Each term of the polynomials consists of the basic integrals expressed in terms of s_a, s_b, p_a and p_b. First we have to evaluate a number of those molecular integrals. Then, using the numerical values of the integrals, we can express $E(R)$ for each fixed value of

* Length: $a_0 = 0.5292 \times 10^{-8}$cm (Bohr radius).
Energy: $e^2/a_0 = 27.21$ eV (twice the ionization energy of hydrogen).

R as a function of λ_g and λ_u. The next and final task is to minimize the energy $E(R)$ with respect to the parameters λ_g and λ_u at each value of R. The potential energy is calculated by the definition $V(R)=E(R)-E(\infty)$, $E(\infty)$ being –5.6953125 au. The whole procedure entails rather tedious but fairly straightforward exercises. We suppose it is not necessary to describe the details here. An appendix is prepared to present formulas to evaluate the basic integrals needed in the present calculation and the relevant numerical tables.

Table 1. Interatomic potential for He_2.

R (Å)	λ_g	λ_u	$V(R)$ (au)	$V(R)$ (au) ($\lambda_g=\lambda_u=0$)	Slater-Kirkwood
0.6272	0.010	−0.225	7.958×10^{-1}	9.137×10^{-1}	
0.9407	0.015	−0.085	2.096×10^{-1}	2.283×10^{-1}	
1.254	0.010	−0.040	5.309×10^{-2}	5.790×10^{-2}	
1.568	0.010	−0.0245	1.232×10^{-2}	1.417×10^{-2}	
1.881	0.0085	−0.0150	2.507×10^{-3}	3.312×10^{-3}	2.314×10^{-3}
2.195	0.0060	−0.0075	4.012×10^{-4}	7.333×10^{-4}	4.222×10^{-4}
2.509	0.0040	−0.0045	3.644×10^{-5}	1.550×10^{-4}	3.466×10^{-5}
3.110	*0.0014	*−0.0014	*-2.14×10^{-5}		-2.696×10^{-5}

*Values due to Griffing and Wehner[1].

The results of the present calculation are shown in Table 1. In the actual calculation the dependence of $E(R)$ on λ_g and λ_u at each distance has been examined, as a rule, at intervals of 0.5 in significant figures. It is observed that λ_g is positive and λ_u is negative consistently for the whole range of the internuclear distance. This implies that the orbital φ_g has a tendency of contracting into the inner region and the orbital φ_u wants to extend itself to the outer region. This is just what we have expected. Repulsive energies of 21.65 eV (0.7958 au) at $R=0.6272$Å and 5.702eV (0.2096 au) at $R=0.9407$Å are favorably compared with the corresponding values of Griffing and Wehner ($\lambda_g=\lambda_u=0$), 24.86eV (0.9137 au) and 6.212 eV (0.2282 au), where they failed to improve the values by varying the effective charge parameter Z. The last column of Table 1 gives values from the so-called Slater-Kirkwood potential:

$$[770\exp(-4.60R)-(1.49/R^6)]\times10^{-12}\,\text{erg.}\quad (R\,\text{in Å}) \tag{16}$$

This is a typical example of "theoretically calculated" potentials and has been widely used. So far, the short range repulsive potential and the attractive dispersion energy are calculated separately and then added up to describe the whole curve of the potential energy between two normal helium atoms. Approximations used in the calculation of the above inverse six-power dispersion energy are not valid for shorter internuclear distances. In the present treatment such an inconsistency is eliminated, and it is the most gratifying feature of our calculation. It is of interest to see that in the region of $2\text{Å}\lesssim R\lesssim3\text{Å}$ the agreement between the Slater-Kirkwood potential and ours is rather excellent. Of course, this does not necessarily mean that the calculated energy values are close enough to the actual ones, but at least, we may say that our calculation gives

values not far from the real truth. It is instructive to note that the same kind of deformations of the molecular orbitals in the single configuration approximation can cover rather effectively the whole range of interaction between two normal helium atoms, from the repulsive region to the van der Waals region in tolerable accuracy.

Appendix

Evaluation formulas and numerical values are given for two-center electron repulsion integrals of hybrid and exchange types used in this treatment. Coulomb integrals are omitted because we can find the relevants formulas and numerical values in Roothaan's works[9)10)]. The integrals can be expressed generally in the form,

$$[\Omega(1)|\bar{\Omega}(2)]=\iint \Omega(1)\,(1/r_{12})\,\bar{\Omega}(2)\,d\tau_1 d\tau_2,$$

where Ω and $\bar{\Omega}$ are charge distributions consisting of products of two atomic orbitals. For example,

$$[s_a s_a|s_a s_b]=\iint s_a(1)\,s_a(1)\,(1/r_{12})\,s_a(2)\,s_b(2)\,d\tau_1 d\tau_2,$$

(hydrid)

$$[s_a s_b|s_a s_b]=\iint s_a(1)\,s_b(1)\,(1/r_{12})\,s_a(2)\,s_b(2)\,d\tau_1 d\tau_2,$$

(exchange)

and others may be inferred according to the above examples. In the following formulas, $A_n(\alpha)$, $B_n(\alpha)$, $G_\tau^\nu(m,\alpha)$, $W_\tau^\nu(m,n;\alpha)$ are the functions defined and tabulated in KAIK's table[11)], namely,

$$A_n(\alpha)=\int_1^\infty e^{-\alpha\xi}\xi^n d\xi,$$

$$B_n(\alpha)=\int_{-1}^{+1} e^{-\alpha\eta}\eta^n d\eta,$$

$$G_\tau^\nu(m,\alpha)=\int_{-1}^{+1} e^{-\alpha\eta}P_\tau^\nu(\eta)\,\eta^m(1-\eta^2)^{\nu/2}d\eta,$$

$$W_\tau^\nu(m,n;\alpha)=\int_1^\infty\int_1^\infty Q_\tau^\nu(\xi_>)\,P_\tau^\nu(\xi_<)\,e^{-\alpha(\xi_1+\xi_2)}\xi_1^m\xi_2^n(\xi_1^2-1)^{\nu/2}(\xi_2^2-1)^{\nu/2}d\xi_1\,d\xi_2,$$

where P_τ^ν and Q_τ^ν are the associated Legendre functions of the 1st and 2nd kinds, respectively, and $\xi_>$ is the larger of ξ_1 and ξ_2 and $\xi_<$ the smaller.

(I) *Hybrid integrals*

$$[s_a s_a|s_a s_b]=Z\frac{\alpha^2}{4}[4A_1(\alpha)+2\{A_0(2\alpha)B_1(\alpha)-A_1(2\alpha)B_0(\alpha)\}$$

$$+\alpha\{A_0(2\alpha)B_2(\alpha)-A_2(2\alpha)B_0(\alpha)\}]$$

$$\left.\begin{matrix}[s_a s_a|s_a p_b] \\ [s_a s_a|p_a s_b]\end{matrix}\right\} = Z\frac{\alpha^6}{120}[\{15W_0(2,2)-5W_0(0,2)\}G_0(0)$$

$$-\{15W_0(0,2)-5W_0(0,0)\}G_0(2)$$
$$\pm\{9W_1(1,2)-15W_1(2,3)\}G_1(0)$$
$$\mp\{9W_1(0,1)-15W_1(0,3)\}G_1(2)$$
$$-10W_2(0,2)G_2(0)+10W_2(0,0)G_2(2)$$
$$\pm 6W_3(1,2)G_3(0)\mp 6W_3(0,1)G_3(2)]$$

$$[s_a p_a|s_a s_b]=Z\frac{\alpha^6}{24}[\{3W_0(2,2)-W_0(0,2)\}G_0(0)$$

$$+\{3W_0(2,3)-W_0(0,3)\}G_0(1)$$
$$-\{3W_0(0,2)-W_0(0,0)\}G_0(2)$$
$$-\{3W_0(1,2)-W_0(0,1)\}G_0(3)$$
$$-2\{W_2(0,2)G_2(0)+W_2(0,3)G_2(1)$$
$$-W_2(0,0)G_2(2)-W_2(0,1)G_2(3)\}]$$

$$[s_a s_a|p_a p_b]=Z\frac{\alpha^7}{1680}[7\{-5W_0(0,2)+18W_0(2,2)-5W_0(2,4)\}G_0(0)$$

$$-7\{-5W_0(0,0)+18W_0(0,2)-5W_0(0,4)\}G_0(2)$$
$$+10\{-7W_2(0,2)+6W_2(2,2)-7W_2(2,4)\}G_2(0)$$
$$-10\{-7W_2(0,0)+6W_2(0,2)-7W_2(0,4)\}G_2(2)$$
$$+24W_4(2,2)G_4(0)-24W_4(0,2)G_4(2)]$$

$$\left.\begin{matrix}[s_a p_a|s_a p_b] \\ [s_a p_a|p_a s_b]\end{matrix}\right\} = Z\frac{\alpha^7}{240}(5[\{3W_0(2,2)-W_0(0,2)\}G_0(0)$$

$$+\{3W_0(2,3)-W_0(0,3)\}G_0(1)$$
$$-\{3W_0(0,2)-W_0(0,0)\}G_0(2)$$
$$-\{3W_0(1,2)-W_0(0,1)\}G_0(3)]$$
$$\pm 3[\{3W_1(1,2)-5W_1(2,3)\}G_1(0)$$
$$+\{3W_1(1,3)-5W_1(3,3)\}G_1(1)$$
$$-\{3W_1(0,1)-5W_1(0,3)\}G_1(2)$$
$$-\{3W_1(1,1)-5W_1(1,3)\}G_1(3)]$$
$$-10[W_2(0,2)G_2(0)+W_2(0,3)G_2(1)$$
$$-W_2(0,0)G_2(2)-W_2(0,1)G_2(3)]$$

$$\pm 6[W_3(1,2)G_3(0)+W_3(1,3)G_3(1)$$
$$-W_3(0,1)G_3(2)-W_3(1,1)G_3(3)])$$

$$[p_\alpha p_\alpha|s_\alpha s_b]=Z\frac{\alpha^7}{48}[\{3W_0(2,2)-W_0(0,2)\}\{G_0(0)-G_0(4)\}$$
$$+\{6W_0(2,3)-2W_0(0,3)\}G_0(1)$$
$$+\{3W_0(2,4)-W_0(0,4)-3W_0(0,2)+W_0(0,0)\}G_0(2)$$
$$-\{6W_0(1,2)-2W_0(0,1)\}G_0(3)]$$
$$-2W_2(0,2)\{G_2(0)-G_2(4)\}-4W_2(0,3)G_2(1)$$
$$-2\{W_2(0,4)-W_2(0,0)\}G_2(2)+4W_2(0,1)G_2(3)]$$

$$[s_\alpha p_\alpha|p_\alpha p_b]=Z\frac{\alpha^8}{3360}\Big(7[\{-5W_0(0,2)+18W_0(2,2)-5W_0(2,4)\}G_0(0)$$
$$+\{-5W_0(0,3)+18W_0(2,3)-5W_0(3,4)\}G_0(1)$$
$$-\{-5W_0(0,0)+18W_0(0,2)-5W_0(0,4)\}G_0(2)$$
$$-\{-5W_0(0,1)+18W_0(1,2)-5W_0(1,4)\}G_0(3)]$$
$$+10[\{-7W_2(0,2)+6W_2(2,2)-7W_2(2,4)\}G_2(0)$$
$$+\{-7W_2(0,3)+6W_2(2,3)-7W_2(3,4)\}G_2(1)$$
$$-\{-7W_2(0,0)+6W_2(0,2)-7W_2(0,4)\}G_2(2)$$
$$-\{-7W_2(0,1)+6W_2(1,2)-7W_2(1,4)\}G_2(3)]$$
$$+24[W_4(2,2)G_4(0)+W_4(2,3)G_4(1)$$
$$-W_4(0,2)G_4(2)-W_4(1,2)G_4(3)]\Big)$$

$$\left.\begin{matrix}[p_\alpha p_\alpha|s_\alpha p_b]\\ [p_\alpha p_\alpha|p_\alpha s_b]\end{matrix}\right\}=Z\frac{\alpha^8}{480}\Big(5[\{3W_0(2,2)-W_0(0,2)\}\{G_0(0)-G_0(4)\}$$
$$+\{6W_0(2,3)-2W_0(0,3)\}G_0(1)$$
$$+\{3W_0(2,4)-W_0(0,4)-3W_0(0,2)+W_0(0,0)\}G_0(2)$$
$$-\{6W_0(1,2)-2W_0(0,1)\}G_0(3)]$$
$$\pm 3[\{3W_1(1,2)-5W_1(2,3)\}\{G_1(0)-G_1(4)\}$$
$$+\{6W_1(1,3)-10W_1(3,3)\}G_1(1)$$
$$+\{-5W_1(3,4)+3W_1(1,4)+5W_1(0,3)-3W_1(0,1)\}G_1(2)$$
$$-\{6W_1(1,1)-10W_1(1,3)\}G_1(3)]$$
$$-10[W_2(0,2)\{G_2(0)-G_2(4)\}$$

$$+2W_2(0,3)G_2(1)+\{W_2(0,4)-W_2(0,0)\}G_2(2)$$
$$-2W_2(0,1)G_2(3)]$$
$$\pm 6[W_3(1,2)\{G_3(0)-G_3(4)\}$$
$$+2W_3(1,3)G_3(1)+\{W_3(1,4)-W_3(0,1)\}G_3(2)$$
$$-2W_3(1,1)G_3(3)]\Big)$$

$$[p_a p_a|p_a p_b]=Z\frac{\alpha^9}{6720}\Big(7[\{-5W_0(0,2)+18W_0(2,2)-5W_0(2,4)\}$$
$$\times\{G_0(0)-G_0(4)\}$$
$$+\{-10W_0(0,3)+36W_0(2,3)-10W_0(3,4)\}G_0(1)$$
$$+\{5W_0(0,0)-18W_0(0,2)+18W_0(2,4)$$
$$-5W_0(4,4)\}G_0(2)-\{-10W_0(0,1)+36W_0(1,2)$$
$$-10W_0(1,4)\}G_0(3)]$$
$$+10[\{-7W_2(0,2)+6W_2(2,2)-7W_2(2,4)\}\{G_2(0)-G_2(4)\}$$
$$+\{-14W_2(0,3)+12W_2(2,3)-14W_2(3,4)\}G_2(1)$$
$$+\{7W_2(0,0)-6W_2(0,2)+6W_2(2,4)-7W_2(4,4)\}G_2(2)$$
$$-\{-14W_2(0,1)+12W_2(1,2)-14W_2(1,4)\}G_2(3)]$$
$$+24[W_4(2,2)\{G_4(0)-G_4(4)\}+2W_4(2,3)G_4(1)$$
$$+\{W_4(2,4)-W_4(0,2)\}G_4(2)-2W_4(1,2)G_4(3)]\Big)$$

(II) *Exchange integrals*

$$[s_a s_b|s_a s_b]=Z\frac{\alpha^5}{90}[45W_0(2,2)-30W_0(0,2)+5W_0(0,0)+4W_2(0,0)]$$

$$[s_a s_b|s_a p_b]=Z\frac{\alpha^6}{180}[45W_0(2,2)-30W_0(0,2)+5W_0(0,0)+4W_2(0,0)]$$
$$=\Big(\frac{\alpha}{2}\Big)[s_a s_b|s_a s_b]$$

$$[s_a s_b|p_a p_b]=Z\frac{\alpha^7}{2520}[35W_0(0,0)-231W_0(0,2)+35W_0(0,4)+378W_0(2,2)$$
$$-105W_0(2,4)+28W_2(0,0)-24W_2(0,2)+28W_2(0,4)]$$

$$\left.\begin{matrix}[s_a p_b|s_a p_b] \\ [s_a p_b|p_a s_b]\end{matrix}\right\} = Z\frac{\alpha^7}{8}\left[\frac{1}{9}\{9W_0(2,2) - 6W_0(0,2) + W_0(0,0)\}\right.$$

$$\pm\frac{1}{75}\{9W_1(1,1) - 30W_1(1,3) + 25W_1(3,3)\}$$

$$\left.+\frac{4}{45}W_2(0,0) \pm \frac{4}{175}W_3(1,1)\right]$$

$$[p_a p_b|p_a p_b] = Z\frac{\alpha^9}{32}\left[\frac{1}{225}\{25W_0(0,0) - 180W_0(0,2) + 50W_0(0,4)\right.$$

$$+324W_0(2,2) - 180W_0(2,4) + 25W_0(4,4)\}$$

$$+\frac{4}{2205}\{49W_2(0,0) - 84W_2(0,2) + 98W_2(0,4)$$

$$+36W_2(2,2) - 84W_2(2,4) + 49W_2(4,4)\}$$

$$\left.+\frac{64}{11025}W_4(2,2)\right]$$

$$[s_a p_b|p_a p_b] = Z\frac{\alpha^8}{5040}[7\{5W_0(0,0) - 33W_0(0,2) + 5W_0(0,4)$$

$$+54W_0(2,2) - 15W_0(2,4)\}$$

$$-4\{-7W_2(0,0) + 6W_2(0,2) - 7W_2(0,4)\}]$$

$$= \left(\frac{\alpha}{2}\right)[s_a s_b|p_a p_b]$$

Numerical values of the above integrals are given in Tables 2 and 3. Some discrepancies are found between numerical values calculated here and those of Hirschfelder and Linnett[12] in $[p_a p_b|p_a p_b]Z$, $[s_a p_a|s_a p_b]Z$, $[s_a p_b|p_a s_b]Z$, $[s_a s_b|s_a s_b]Z$, $[p_a p_a|p_a p_b]Z$.

Table 2. Hybrid integrals

$\alpha = RZ$	$[s_a s_a\|s_a s_b]/Z$	$[s_a s_a\|s_a p_b]/Z$	$[s_a s_a\|p_a s_b]/Z$	$[s_a p_a\|s_a s_b]/Z$
2	3.0803647×10^{-1}	3.6981733×10^{-1}	2.4625558×10^{-1}	3.8221220×10^{-2}
3	1.6074246×10^{-1}	2.9427873×10^{-1}	1.8794864×10^{-1}	2.4065682×10^{-2}
4	7.6981672×10^{-2}	1.9039060×10^{-1}	1.1753607×10^{-1}	1.2454731×10^{-2}
5	3.4953043×10^{-2}	1.0916325×10^{-1}	6.5601956×10^{-2}	5.8181988×10^{-3}
6	1.5311456×10^{-2}	5.7842934×10^{-2}	3.4025800×10^{-2}	2.5593022×10^{-3}
7	6.5378679×10^{-3}	2.8998003×10^{-2}	1.6767076×10^{-2}	1.0833484×14^{-3}
8	2.7387379×10^{-3}	1.3954049×10^{-2}	7.9558530×10^{-3}	4.4672490×10^{-4}

$\alpha=RZ$	$[s_a s_a \mid p_a p_b]/Z$	$[s_a p_a \mid s_a p_b]/Z$	$[s_a p_a \mid p_a s_b]/Z$	$[p_a p_a \mid s_a s_b]/Z$
2	-3.5351805×10^{-2}	4.2390987×10^{-3}	7.2203341×10^{-2}	2.4571540×10^{-1}
3	1.0797465×10^{-1}	2.1943240×10^{-2}	5.0253806×10^{-2}	1.3698695×10^{-1}
4	1.3846601×10^{-1}	2.0684192×10^{-2}	2.9134731×10^{-2}	6.8659446×10^{-2}
5	1.1410166×10^{-1}	1.4002323×10^{-2}	1.5088712×10^{-2}	3.2140488×10^{-2}
6	7.7585837×10^{-2}	8.0856340×10^{-3}	7.2701795×10^{-3}	1.4375304×10^{-2}
7	4.7110005×10^{-2}	4.2469098×10^{-3}	3.3365288×10^{-3}	6.2283510×10^{-3}
8	2.6508945×10^{-2}	2.0945827×10^{-3}	1.4792165×10^{-3}	2.6367209×10^{-3}

$\alpha=RZ$	$[s_a p_a \mid p_a p_b]/Z$	$[p_a p_a \mid s_a p_b]/Z$	$[p_a p_a \mid p_a s_b]/Z$	$[p_a p_a \mid p_a p_b]/Z$
2	3.2179978×10^{-2}	2.5907865×10^{-1}	2.3235215×10^{-1}	-5.3478619×10^{-2}
3	4.7154886×10^{-2}	2.2350498×10^{-1}	1.8745588×10^{-1}	9.1694848×10^{-2}
4	4.3229975×10^{-2}	1.5377967×10^{-1}	1.2085811×10^{-1}	1.2404144×10^{-1}
5	2.0699942×10^{-2}	9.2239244×10^{-2}	6.8463095×10^{-2}	1.1377397×10^{-1}
6	1.3566483×10^{-2}	5.0521657×10^{-2}	3.5730166×10^{-2}	7.9087660×10^{-2}
7	8.0410639×10^{-3}	2.6191652×10^{-2}	1.7866237×10^{-2}	4.8598624×10^{-2}
8	4.4476073×10^{-3}	1.2947387×10^{-2}	8.5798842×10^{-3}	2.7522644×10^{-2}

Table 3. Exchange integrals

$\alpha=RZ$	$[s_a s_b \mid s_a s_b]/Z$	$[s_a s_b \mid s_a p_b]/Z$	$[s_a s_b \mid p_a p_b]/Z$	$[s_a p_b \mid s_a p_b]/Z$
2	1.8415645×10^{-1}	1.8415645×10^{-1}	-1.1687838×10^{-2}	2.2400267×10^{-1}
3	5.8507960×10^{-2}	8.7761941×10^{-2}	4.7499783×10^{-2}	1.5078940×10^{-1}
4	1.5627203×10^{-2}	3.1254406×10^{-2}	3.2603373×10^{-2}	6.9941051×10^{-2}
5	3.7170295×10^{-3}	9.2925736×10^{-3}	1.3997416×10^{-2}	2.5690740×10^{-2}
6	8.1402316×10^{-4}	2.4420695×10^{-3}	4.7645076×10^{-3}	8.0473282×10^{-3}
7	1.6759955×10^{-4}	5.8659843×10^{-4}	1.3994292×10^{-3}	2.2456471×10^{-3}
8	3.2895900×10^{-5}	1.3158360×10^{-4}	3.7031673×10^{-4}	5.7407367×10^{-4}

$\alpha=RZ$	$[s_a p_b \mid p_a s_b]/Z$	$[p_a p_b \mid p_a p_b]/Z$	$[s_a p_b \mid p_a p_b]/Z$
2	1.4431023×10^{-1}	4.3991508×10^{-2}	-1.1687838×10^{-2}
3	1.1249642×10^{-1}	6.1409292×10^{-2}	7.1249674×10^{-2}
4	5.5076574×10^{-2}	7.8505099×10^{-2}	6.5206745×10^{-2}
5	2.0772128×10^{-2}	5.6941825×10^{-2}	3.4993540×10^{-2}
6	6.6050886×10^{-3}	2.9409874×10^{-2}	1.4293523×10^{-2}
7	1.8605419×10^{-3}	1.2182024×10^{-2}	4.8980023×10^{-3}
8	4.7859514×10^{-4}	4.3179227×10^{-3}	1.4812669×10^{-3}

References

1) V.Griffing and J. F. Wehner, J. Chem. Phys. **23** (1955), 1024.
2) M. Sakamoto and E. Ishiguro, Prog. Theor. Phys. **15** (1956), 37.
3) T. Inui, Proc. Phys-Math. Soc. Japan **20** (1938), 770; ibid. **23** (1949), 992.

4) A. Nordsieck, Phys. Rev. **58** (1940), 310.
5) S. Huzinaga, Prog. Theor. Phys. **18** (1957), 139.
6) C. A. Coulson, Trans. Farad. Soc. **33** (1937), 1479.
7) N. Rosen, Phys. Rev. **41** (1932), 255.
8) B. N. Dickinson, J. Chem. Phys. **1** (1933), 317.
9) C. C. J. Roothaan, J. Chem. Phys. **19** (1951), 1445.
10) C. C. J. Roothaan, Special Technical Report, Laboratory of Molecular Structure and Spectra, Department of Physics, The University of Chicago (1955).
11) M. Kotani, A. Amemiya, E. Ishiguro and T. Kimura, *Table of Molecular Integrals*, Maruzen Co., Ltd., Tokyo, Japan (1955).
12) J. O. Hirschfelder and J. W. Linnett, J. Chem. Phys. **18** (1950), 130.

Progress of Theoretical Physics, Vol. 20, No. 5, November 1958

The First Forbidden β-Decay*

Tsuneyuki KOTANI** and Marc ROSS

Department of Physics, Indiana University
Bloomington, Indiana

(Received July 6, 1958)

The energy spectrum, longitudinal polarization, and various β-γ correlations, with and without measurement of β or γ-ray polarizations, are calculated for the first forbidden β-decay. The theoretical expressions include finite nuclear size effects, in the form of a simple modification of older expressions. A short discussion of relevant Coulomb functions is presented. Interference from so-called third forbidden transitions is also included. The expressions are presented in a reasonably accurate expansion such that the energy dependence, dependence on real and imaginary parts of coupling constants, and dependence on independent nuclear matrix elements, are clearly revealed. Possibilities for testing the time reversal invariance of the interaction are discussed. The β-γ directional correlation is probably most useful for this purpose among the various possibilities in the first forbidden transition. The transverse β polarization is probably not useful for this purpose (it may, however, be used to determine nuclear matrix elements). Determination of other general properties of the β-decay interaction from first forbidden decays is probably even more difficult. Special information that might, however, be obtained from examination of a 0-0 transition is discussed. It is proposed that there is a fruitful field for determinig properties of specific nuclei by measurement of several properties of a single decay. Both accurate and simplified expressions are presented to show what may be looked for. Absolute measurement of γ polarization following β-decay and measurement of the β-γ correlation and its energy dependence are very promising experiments; as are of course, (the widely known) accurate spectrum and longitudinal polarization measurements. A special discussion is presented of the RaE case, where it is pointed out that accurate longitudinal polarization measurements could be very helpful.

§ 1. Introduction

After Lee and Yang raised the question of violation of invariance of weak interactions under space inversion, charge conjugation, and time reversal,[1] Wu and her co-workers, and many others, confirmed that space inversion and charge conjugation invariances are violated in β decay[2,3]. One consequence is that there are more possible experiments; for example, those involving polarization measurements of the β-rays. Theoretical formulae for various phenomena have been given by many authors. However, in the case of first forbidden transitions, this work is not sufficiently complete for analysis of many experimental results, e.g. expansions of the transition probability expressions are sometimes cut-off too soon and nuclear finite size effects are not taken into account very accurately. The

* Supported in part by the National Science Foundation.
** On leave from Tokyo Metropolitan University, Setagaya, Tokyo.

purpose of this paper is to discuss classes of possible experiments in first forbidden β decay, those involving measurements on the β ray and a possible following γ ray, together with theoretical predictions which will allow an accurate analysis of the experiments. We will also discuss, with the aid of some simplified expressions, which experiments would be most likely to yield fruitful information.

At present there are several fundamental questions about $\beta^{\mp}$ decay which are independent of the individual nucleus involved. We will discuss the following: 1) Which interactions among the types S, V, T, A and P are necessary? 2) Is the β-decay interaction invariant under time reversal? 3) What are the relative signs and magnitudes of the coupling constants?

There are other fundamental questions, such as: 4) Is there a lepton conservation law in β-decay? 5) Is there a fundamental difference in the role of β^- and β^+ fields in the interaction?[4] 6) Is there a non local property in the β-decay interaction?[5] We will not discuss these questions.

In addition to the above questions, violation of parity conservation gives rise to the interesting possibility that a great deal may be learned about particular nuclear matrix elements. The various experiments, which we now know are possible, will give independent information on the nuclear matrix elements involved. We will discuss this possibility in the following sections.

One other important question can now be considered to be answered, i.e. the relation between the parity conserving and non-conserving parts of the β decay interaction (we use the convenient expression "parity non-conserving interaction" although the observation of parity non-conservation arises from an interference effect involving these parts). The answer has been given by measurements of longitudinal polarization of β-ray (as will be discussed in § 2). At present we will just state as an assumption the implication of this result:

Assumption I. *There is no interference between (S, T, P) and (V, A) combinations of interactions. The (S, T, P) and (V, A) combinations can thus be treated separately.*

Concerning the question of how much (S, T) and how much (V, A) interactions are present, we have fairly convincing new information that (V, A) is present and not (S, T). Consider first the electron-neutrino correlation experiment[6]. The neutron and Ne^{19} correlation coefficients show that we have either (S, T) and or (V, A) (but not (S, A) or (T, V) combinations). The A^{35} results show that the dominant Fermi type interaction is V. In principle the He^6 correlation can decide between the Gamov-Teller interactions T and A; but at present the He^6 experiments do not yield definite results.[7] Meanwhile Goldhaber, Grodzins and Sunyar[8] have observed the relation between the neutrino spin and its momentum direction (i.e. its helicity) in the Eu^{152m} decay. They find that A must be the dominant G-T interaction. It is still of some interest to point out how to distinguish between (S, T) and (V, A). The characteristic difference between (S, T, P) and (V, A) combinations is that the operators in the scalar, tensor, and

pseudoscalar interactions, $O_S=\beta$, $O_T=(\beta\sigma, \beta\alpha)$ and $O_P=\beta\gamma_5$, include one more β-matrix than each corresponding vector and pseudovector operator, $O_V=(1, i\alpha)$ and $O_A=(\sigma, i\gamma_5)$. In calculating the total transition probability we can sum over the (unobserved) neutrino states. This sum can be made by using the so-called Casimir trick (projection operator). The transition probability will then contain the factor

$$O_j\left\{\frac{q+\boldsymbol{\alpha}\cdot\boldsymbol{q}}{2q}\,e^{i\boldsymbol{q}\boldsymbol{r}}\right\}O_k^+ . \qquad (1\cdot1)$$

Here q and $\boldsymbol{q}$ are the neutrino energy and momentum, and $\boldsymbol{r}$ the coordinate vector (which will be integrated in a nuclear matrix element). For parity conserving processes it is then clear that the difference between the (S, T, and P) and (V, A) combinations will be reflected only in a change in sign of the $\boldsymbol{\alpha}\cdot\boldsymbol{q}$ term. The same result follows for the parity non-conserving processes if the different relations (as given in Sec. 2, eq. (2·10)) between parity conserving and non-conserving parts in the (S, T, P) and (V, A) combinations are taken into account. Therefore in allowed transitions, where we may assume $\exp(i\boldsymbol{q}\cdot\boldsymbol{r})=1$, it is necessary to observe the neutrino direction in order to distinguish (S, T) from (V, A). In forbidden transitions, expanding $\exp(i\boldsymbol{q}\cdot\boldsymbol{r})$, it is seen from expression (1·1) that, even if we don't observe the neutrino direction directly, there will be terms with odd powers of q which have different signs for the (S, T, P) and (V, A) combinations. Unfortunately, in the first forbidden transitions, coulomb correction terms (from the β-particle wave function) which are independent of q are generally much larger than any terms involving, say, q to the first power. Thus we again may not expect to distinguish reliably between (S, T, P) and (V, A) without measuring the nuclear recoil. In this paper we will not consider in full detail experiments including the polarized nucleus or measurement of the nuclear recoil. We will discuss them briefly in § 5.

The problem of whether or not the pseudoscalar interaction contributes significantly in β decay cannot be determined from allowed transitions where there is no contribution from this type of interaction. We will discuss in § 3 how the P contribution might be estimated in the first forbidden decay. Actually there are good theoretical reasons for believing that the nuclear matrix element from the P interaction is very much smaller than the typical nuclear matrix element in the 1st forbidden β decay.[9]

Before considering the question of invariance under time reversal, let us first make some general observations about experiments involving measurement of direction and polarization of the β particle and a possible following γ-ray. These experiments can be symbolically represented in terms of the unit vectors $\boldsymbol{p}$, $\boldsymbol{k}$, $\boldsymbol{s}$, $\boldsymbol{t}$, where $\boldsymbol{p}$ and $\boldsymbol{k}$ are electron and photon directions, $\boldsymbol{s}$ is the electron polarization and $\boldsymbol{t}$ the direction of circular polarization of the γ-ray. We note that $\boldsymbol{t}$ can only occur in the combination $(\boldsymbol{k}\cdot\boldsymbol{t})$. Further, as a result of parity conservation in the electromagnetic interaction, $\boldsymbol{k}$ must occur to an even power. For this reason the factor $(\boldsymbol{k}\cdot\boldsymbol{t})$ is usually not explicitly shown; it being taken for granted that the circular polarization must be measured if $\boldsymbol{k}$ occurs an odd number of times in the rest of the distribution. We will adopt this convention

in the remaining sections of the paper. Of course, the quantities $(\boldsymbol{k}\cdot\boldsymbol{t})$ and $\boldsymbol{s}$ need only be considered to zero power and linearly. It will be shown in § 2 that while $\boldsymbol{p}$ may occur at most to second power in allowed transitions, it may occur to fourth power in first forbidden transitions.

In Table 1 the possible experiments are shown along with their oddness or evenness under the two transformations :

$$P: \quad (\boldsymbol{p}, \boldsymbol{s}, \boldsymbol{k}, \boldsymbol{t}) \rightarrow (-\boldsymbol{p}, \boldsymbol{s}, -\boldsymbol{k}, \boldsymbol{t}), \tag{1·2}$$

and

$$T: \quad (\boldsymbol{p}, \boldsymbol{s}, \boldsymbol{k}, \boldsymbol{t}) \rightarrow (-\boldsymbol{p}, -\boldsymbol{s}, -\boldsymbol{k}, -\boldsymbol{t}). \tag{1·3}$$

We note in Table 1 the possibilities of observing :

A : longitudinal polarization of β-rays,

B : β-γ directional correlation,

D : transverse β-ray polarization in the plane of the β and γ-rays,

E : transverse polarization perpendicular to the plane of β- and γ-rays.

Certain distributions involve the measurement circularly polarized γ-rays. (if the polarization is not observed, these terms disappear.)

Table 1

	Experiment*	Transformation P	Transformation T
A	$(\boldsymbol{p}\cdot\boldsymbol{s})\ P_n(\boldsymbol{p}\cdot\boldsymbol{k})$	odd	even
	$(\boldsymbol{p}\cdot\boldsymbol{s})\ P_{n'}(\boldsymbol{p}\cdot\boldsymbol{k})\ (\boldsymbol{k}\cdot\boldsymbol{t})$	even	even
B	$P_n(\boldsymbol{p}\cdot\boldsymbol{k})$	even	even
	$P_{n'}(\boldsymbol{p}\cdot\boldsymbol{k})\ (\boldsymbol{k}\cdot\boldsymbol{t})$	odd	even
D	$((\boldsymbol{p}\times(\boldsymbol{p}\times\boldsymbol{k}))\cdot\boldsymbol{s})\ P_{n'}(\boldsymbol{p}\cdot\boldsymbol{k})$	odd	even
	$((\boldsymbol{p}\times(\boldsymbol{p}\times\boldsymbol{k}))\cdot\boldsymbol{s})\ P_n(\boldsymbol{p}\cdot\boldsymbol{k})\ (\boldsymbol{k}\cdot\boldsymbol{t})$	odd	even
E	$(\boldsymbol{s}\cdot\boldsymbol{p}\times\boldsymbol{k})\ P_{n'}(\boldsymbol{p}\cdot\boldsymbol{k})$	even	odd
	$(\boldsymbol{s}\cdot\boldsymbol{p}\times\boldsymbol{k})\ P_n(\boldsymbol{p}\cdot\boldsymbol{k})\ (\boldsymbol{k}\cdot\boldsymbol{t})$	odd	odd

* In every case n is even and n' is odd and P_n is a Legendre function.

At this point we should mention the possible experiments involving a β and two following γ-rays. One finds that $(\boldsymbol{p}\cdot\boldsymbol{k}_1\times\boldsymbol{k}_2)^{n'}\ P_{l'}(\boldsymbol{k}_1\cdot\boldsymbol{k}_2)$, with n' and l' odd, which is odd with respect to P and T, vanishes (provided strong interactions are invariant with respect to time reversal). The triple correlation $(\boldsymbol{p}\cdot\boldsymbol{k}_1\times\boldsymbol{k}_2)^{n}\ P_l(\boldsymbol{k}_1\cdot\boldsymbol{k}_2)$, with n and l even, does not vanish, but it is more difficult to measure this term than other distributions which yield the same information.

In order to test the time reversal assumption in β-decay, it is necessary to determine whether the imaginary parts of the coupling constants vanish.[10] It is recalled that in the case of parity non-conservation it was sufficient to detect a distribution odd with respect to the transformation P, (1·2). It is not sufficient, with time reversal, to detect a distribution odd with respect to the transformation T, (1·3). Generally if Coulomb (or other interaction) corrections to the outgoing wavefunctions are taken into account,

the situation is more complicated. Consider observing a distribution O_+ or O_- where $+$ and $-$ mean even and odd under the transformation T, respectively. Then the first Coulomb correction to these observables (using, as an example, the usual β-decay coupling constants C_j) yields the result:

$$\langle O_+ \rangle = \sum_{j,k} \{a_{jk}^{(+)} \mathrm{Re}\, C_j C_k^* + \alpha Z b_{jk}^{(+)} \mathrm{Im}\, C_j C_k^*\}, \tag{1·4}$$

and

$$\langle O_- \rangle = \sum_{j,k} \{a_{jk}^{(-)} \mathrm{Im}\, C_j C_k^* + \alpha Z b_{jk}^{(-)} \mathrm{Re}\, C_j C_k^*\}, \tag{1·5}$$

where a_{jk} and b_{jk} are numerical factors, including the nuclear matrix elements and depending on the coupling types (j and k), and αZ is the fine structure constant times the nuclear charge. It is seen that measurement of a distribution O_- might usually be expected to yield more definite information on time reversal invariance than O_+ if a_{jk} is larger than b_{jk}. The odd observables O_- will generally involve triple scalar products like Type E in Table 1. In considering experiments involving oriented nuclei with orientation $\boldsymbol{J}$, Morita[11], Curtis and Lewis[12], and Dolginov[12] gave the theoretical expression for $(\boldsymbol{J}\cdot\boldsymbol{p}\times\boldsymbol{k})\,(\boldsymbol{J}\cdot\boldsymbol{k})^{n'}$, where n' is an odd integer. There is no contribution from $b^{(-)}$ of (1·5) in the allowed transition case. Ambler et al.[13] did this experiment with Co^{58} and Mn^{52}. Their results are consistent with no violation of time reversal invariance. It is not yet established, however, that the extent of time reversal violation is really small. Jackson, Treiman and Wyld[14] proposed the measurement of $(\boldsymbol{J}\cdot\boldsymbol{p}\times\boldsymbol{q})$ to test time reversal, where again, $b^{(-)}=0$ in the allowed transition. Experimental results (using a polarized neutron beam and measuring the recoil proton) for this case are not yet available*. The last one is $(\boldsymbol{k}\cdot\boldsymbol{p}\times\boldsymbol{q})$,[11),15)] where the γ-ray should be measured by the method of the nuclear fluorescence scattering to observe the nuclear recoil. Other possible triple scalar products will not be useful in allowed transition, if there is no $(S, T)-(V, A)$ interference, because $a^{(-)}$ of eq. (1·5) is essentially zero.

Let us consider odd observables in first forbidden decays. We will find (in § 4) that both transverse polarization experiments of type E (Table 1) have the drawback that $a^{(-)}$ is generally smaller than, or of the same order as, $b^{(-)}$ (in 1·5).

Even in the case that we do not measure an odd observable, we can test time reversal by comparing the real part of coupling constants with their absolute value. In allowed transitions, we can do this comparison by observing the interference term between Fermi and Gamow-Teller interactions, for example, in the even distributions (O_+) : $(\boldsymbol{J}\cdot\boldsymbol{p})$ or $(\boldsymbol{p}\cdot\boldsymbol{k})$.[16] This comparison of the real parts of coupling constants with their absolute values depends to some extent on knowledge of the nuclear matrix element, except for the electron angular distribution from polarized neutrons. The recent experimental results of Telegdi et al.[17] show that the contribution from the imaginary part is fairly small. The remaining way to test time reversal is to measure, for example, $b_{jk}^{(+)}\,\mathrm{Im}\,C_jC_k^*$ in eq. (1·4) directly, as they occur in an even distribution, O_+. For this purpose, the energy dependence of b_{jk} must be different from that of a_{jk}. Both possibilities for observing,

* **Note added in proof :** See the paper by Clark, Robson and Nathans, Phys. Rev. Letter, **1** (1958), 100.

for example, $a_{jk}\,\mathrm{Re}\,C_j C_k^*$ and $b_{jk}\,\mathrm{Im}\,C_j C_k^*$ terms in eq. (1·4), will be discussed in § 3 and 4.

In order to determine the relative signs or the magnitude of the coupling constants C_A and C_T, say, the relative signs and magnitudes of the nuclear matrix elements have to be known. For this reason the neutron decay is of special interest. In any other special cases where we have some information about the matrix elements, we can also deduce relations among the coupling constants.

In § 2, the general expression for the probability of a β-γ transition will be given. In § 3, the energy spectrum shape and longitudinal polarization will be discussed for various types of the 1st forbidden transition, (such as, the 1–0 transition, e.g. RaE, etc.), with an examination of accurate expressions for these processes. In § 4, the leading or largest coefficients of each type of observable distribution will be given, so that qualitative discussion is easy. Appendix A contains a definition of the first forbidden transition and discussion of "third forbidden contributions." In Appendix B we illustrate the modification of the formal expression for the transition probability, in order to take into account finite nuclear size effects. Appendix C contains the detailed theoretical results for various quantities.

§ 2. General formula for first forbidden transitions

Pauli[18] has shown that, if we do not assume lepton conservation, there will in general be twenty complex coupling constants in the β-interaction. We define these coupling constants by writing the interaction Hamiltonian density as:

$$H_{\mathrm{int}}=\sum_j (\psi_p^+ O_j \psi_n)\,(\psi_e^+ O_j[(C_j+C_j'\gamma_5)\psi_\nu+(D_j+D_j'\gamma_5)\psi_\nu^C]) \qquad (2\cdot1)$$
$$+\mathrm{h.c.}$$

Here the wave functions ψ_p and ψ_n represent the final and initial nucleon states in the β^- decay, and vice versa in the β^+ decay. The index j runs over the usual interaction terms S, V, T, A, and P. The charge conjugate field ψ^C is defined, for example, by Pauli[18]. The notation C_j, C_j' and O_j is that used by Lee and Yang[*]. The constants D_j and D_j' are defined by Kahana and Pursey[19].

Several authors have shown[15, 20] that if we do not consider the double processes, for example, the double β-decay and the "chain reaction" of the neutrino, the coupling constants appear in just two kinds of combinations in single β-decay, i.e.:

$$K_{jk}=C_jC_k^*+C_j'C_k'^*+D_jD_k^*+D_j'D_k'^*, \qquad (2\cdot2)$$

$$L_{jk}=C_jC_k'^*+C_j'C_k^*+D_jD_k'^*+D_j'D_k^*. \qquad (2\cdot3)$$

We will show some theoretical results using this notation in Appendix B. If the theory

[*] O_j stands for $O_S=\gamma_4$, $O_V=\gamma_4\gamma_\mu$, $O_T=\gamma_4(\gamma_\mu\gamma_\nu-\gamma_\nu\gamma_\mu)(-i/2\sqrt{2})$, $O_A=i\gamma_4\gamma_\mu\gamma_5$ and $O_P=\gamma_4\gamma_5$ where the γ matrices are defined: $\gamma_k=-i\beta\alpha_k$, $\gamma_4=\beta$ $\gamma_5=\gamma_1\gamma_2\gamma_3\gamma_4$ and $\alpha_k=-\gamma_5\sigma_k$. Throughout the paper we will use the units to $\hbar=c=m_e=1$, where m_e is the electron mass.

is invariant under space inversion, P, then

$$L_{jk}=0. \tag{2·4}$$

On the other hand invariance with respect to time reversal, *i.e.* the transformation T, requires

$$\text{Im } K_{jk}=\text{Im } L_{jk}=0\,,$$

or

$$\text{Im } C_j=\text{Im } C_j'=0\,,\ \text{etc.} \tag{2·5}$$

It is convenient to consider a certain simplification, at this point, which can be deduced from the longitudinal polarization of β-rays. The longitudinal polarization is defined as

$$P_L=\frac{N(+)-N(-)}{N(+)+N(-)}\,, \tag{2·6}$$

where $N(+)$ and $N(-)$ are the numbers of β-rays with their spins parallel and antiparallel to their momenta, respectively. The experimental result for P_L is[21]

$$P_L\approx\mp\,(p/W) \tag{2·6′}$$

for electron (e^-) and positron (e^+), respectively.

Let us consider two, say positive energy, states of form $(1-\gamma_5)\psi_\nu$ with $(C_j=-C_j')$ and $(1+\gamma_5)\psi_\nu$ with $(C_j=C_j')$, as predicted by the two component neutrino theory. The spin of the neutrino (ν) is exactly parallel and antiparallel to its momentum, respectively, in these states (if the neutrino has zero mass). These are called right (R) and left (L) helicities, respectively. The antineutrino ($\bar{\nu}$) created by these operators, $(1\mp\gamma_5)\psi_\nu$, will, on the other hand, be left and right handed, respectively. If extremely high energy β-rays are considered, where the mass can be considered as negligible, the same considerations apply to the β-rays. Let us consider interaction factors

$$(\psi_e^+O_j(1\pm\gamma_5)\psi_\nu), \tag{2·7}$$

(compare with eq. (2·1)). We note that this equals

$$(1/2)\,([(1\mp\gamma_5)\psi_e]^+O_j(1\pm\gamma_5)\psi_\nu)\quad\text{for } S,\ T \text{ and } P, \tag{2·8}$$

where O_j is odd in the γ's, and

$$(1/2)\,([(1\pm\gamma_5)\psi_e]^+O_j(1\pm\gamma_5)\psi_\nu)\quad\text{for } V \text{ and } A. \tag{2·9}$$

Therefore, in order to get the left helicity (eq. (2·6′)) for the high energy electron ($e^-(L)$), we find that

$$C_j=-C_j'\quad\text{for}\quad S,\ T,\ P$$

$$C_j=+C_j'\quad\text{for}\quad V,\ A$$

i.e. the lower case of (2·8) and the upper case of (2·9). In other words, for S, T, or P, $e^-(L)$ implies that $e^-(R)$, $\nu(R)$, and $\bar{\nu}(L)$ are produced in the β-decay process,

while for V or A, e^- and ν have the same helicity, so that $e^-(L)$ implies $\nu(L)$, $e^-(R)$ and $\bar{\nu}(R)$. A similar conclusion can be derived in the case with coupling constants D_j and D_j', except for exchanging the name of neutrino (ν) into antineutrino ($\bar{\nu}$) and *vice versa*. (Compare with the last term of H_{int} (2·1)).

Without waiting for more definite experimental confirmation, we will assume for simplicity :

Assumption I'. (Two Component Theory of the Neutrino) :

$$\begin{aligned} &C_j=-C_j' \text{ and } D_j=-D_j' \text{ for } j=S,\ T,\ P, \\ &C_j=+C_j' \text{ and } D_j=D_j' \quad \text{for } j=V,\ A, \end{aligned} \tag{2·10}$$

From this assumption, we see that $K_{jk}=L_{jk}=0$ if $j=S,\ T,\ P$ and $k=V,\ A$, and *vice versa*. Thus, as stated in the introduction, we will consider no $(STP)-(VA)$ interference.

We will further simplify our expressions and discussion by assuming $D_j=D_j'=0$ (except in Appendix B). The expressions are very easily generalized to non-zero D's, however.

Consider now the general β decay process with unoriented nuclei and with no measurement of nuclear recoil. We use the notation $J_0\to(\beta)\to J_1\to(\gamma)\to J_2$ for the decay. The transition probability for β-ray of energy W to $W+dW$, spin direction $\boldsymbol{s}$, momentum direction $\boldsymbol{p}$ to $\boldsymbol{p}+d\boldsymbol{p}$ and γ-ray of direction $\boldsymbol{k}$ to $\boldsymbol{k}-d\boldsymbol{k}$ has the angular properties :

$$N(W,\ \boldsymbol{p},\ \boldsymbol{s},\ \boldsymbol{k},)\,dW\,d\boldsymbol{p}\,d\boldsymbol{k}\sim$$

$$\Lambda_{\nu_1\nu_2}\begin{pmatrix} l_\kappa & l_{\kappa'} & \nu_1 \\ 0 & 0 & 0 \end{pmatrix}\begin{Bmatrix} J_1 & J_1 & n \\ \lambda' & \lambda & J_0 \end{Bmatrix}\begin{Bmatrix} j & j' & n \\ l(\kappa) & l(\kappa') & \nu_1 \\ \frac{1}{2} & \frac{1}{2} & \nu_2 \end{Bmatrix}, \tag{2·11}$$

where the brackets are $3j$, $6j$, and $9j$ symbols,[27] respectively, and

$$\Lambda_{\nu_1\nu_2}=\sum_{\mu_1\mu_2}\begin{pmatrix} \nu_1 & \nu_2 & n \\ \mu_1 & \mu_2 & \mu \end{pmatrix} Y_{\nu_1}^{*\mu_1}(\boldsymbol{p})\,Y_{\nu_2}^{*\mu_2}(\boldsymbol{s})\,Y_n^{\mu}(\boldsymbol{k}). \tag{2·12}$$

Here λ and λ' are the tensor ranks of the operators in the β-decay nuclear matrix elements. The total angular momentum, orbital angular momentum (of the large component), and spin of the β particle are given by $j=|\kappa|-\frac{1}{2}$, $l(\kappa)=|\kappa|-\frac{1}{2}+(\kappa/2|\kappa|)$, and $1/2$, respectively.[28] The index κ runs over all integers, except 0. In allowed transitions $\kappa=\pm1$, so the maximum values of ν_1 and n are $n=1$, $\nu_1=2$. Thus, as mentioned in the introduction, we cannot have a distribution such as $(\boldsymbol{p}\cdot\boldsymbol{k})^2$ and $(\boldsymbol{s}\cdot\boldsymbol{p}\times\boldsymbol{k})(\boldsymbol{p}\cdot\boldsymbol{k})$ in the allowed transition.* In the first forbidden transition the main terms are all associated

* These are not exact statements, since the exact expression for an allowed transition contains contributions from all κ's. The leading terms come from $\kappa=\pm1$. For example, there is a small correction of relative order $\alpha Z p\rho$ (where ρ is the nuclear radius) to the allowed transition, associated with $|\kappa|=2$, which yields the distribution $(\boldsymbol{p}\cdot\boldsymbol{k})^2$ and $(\boldsymbol{s}\cdot\boldsymbol{p}\times\boldsymbol{k})(\boldsymbol{p}\cdot\boldsymbol{k})$.

with $\kappa=\pm1,\pm2$, so the maximum values of n and ν_1 are $n=3$ and $\nu_1=4$.

If γ-ray circular polarization is not measured, the transition probability has the form :

$$N(W,\ \boldsymbol{p},\ \boldsymbol{s},\ \boldsymbol{k})\,dW\,d\boldsymbol{p}\,d\boldsymbol{k}=$$

$$F_0(Z,\ W)\,pW\,(W_0-W)^2 dW\,d\boldsymbol{p}\,d\boldsymbol{k}\,(1+\gamma_1/2)\,(2|\eta_J|^2\textstyle\sum_L|\partial_L|^2)/[(2\pi)^5 2]$$

$$\left.\begin{array}{l} B^{(0)}+B^{(2)}\frac{1}{2}\,(3\,(\boldsymbol{p}\cdot\boldsymbol{k})^2-1) \\ +A^{(0)}\,(\boldsymbol{p}\cdot\boldsymbol{s})+A^{(2)}\,(\boldsymbol{p}\cdot\boldsymbol{s})\,(3\,(\boldsymbol{p}\cdot\boldsymbol{k})^2-1)/2, \\ +D^{(2)}\,(\boldsymbol{s}\cdot\boldsymbol{p}\times(\boldsymbol{k}\times\boldsymbol{p}))\,(\boldsymbol{p}\cdot\boldsymbol{k}) \\ +E^{(2)}\,(\boldsymbol{s}\cdot\boldsymbol{p}\times\boldsymbol{k})\,(\boldsymbol{p}\cdot\boldsymbol{k}) \end{array}\right\}, \qquad (2\cdot13)$$

where η_J will be defined in § 3. (See Table (2) and eqs. (3·6) and (3·7).*)

$$\gamma_\kappa=\sqrt{\kappa^2-(\alpha Z)^2}\,, \qquad (2\cdot14)$$

$$B^{(n)}=\sum_{\lambda\leqq\lambda'}{}' b^{(n)}_{\lambda\lambda'}G_{\lambda\lambda'}(n\,;J_0J_1J_2)\,, \qquad (2\cdot15)$$

and

$$G_{\lambda\lambda'}(n)=(-1)^{J_1-J_0}W(J_1J_1\lambda\lambda'\,;nJ_0)\sqrt{2J_1+1}$$
$$(\textstyle\sum_{LL'}(-1)^{L+L'}F_n(LL'J_2J_1)\,\partial^*_L\partial_{L'}/\sum_L|\partial_L|^2)\,, \qquad (2\cdot16)$$

where

$$F_n(LL'J_2J_1)=(-1)^{J_1-J_2-1}\sqrt{(2J_1+1)\,(2L+1)\,(2L'+1)}$$
$$C(LL'n\,;\,1-1)\,W(J_1J_1LL'\,;nJ_2)\,.$$

The other coefficients $A^{(n)}$, $D^{(n)}$, $E^{(n)}$, are defined as in (2·15) and (2·16) except that we substitute $a^{(n)}_{\lambda\lambda'}$, $d^{(n)}_{\lambda\lambda'}$, $e^{(n)}_{\lambda\lambda'}$, respectively, for the particle parameter $b^{(n)}_{\lambda\lambda'}$. The nuclear matrix elements for the β decay are contained in these particle parameters. The calculated results for these particle parameters are given in Appendix B and C**. The ∂_L's are reduced matrix elements for the 2^L pole γ-ray emission (if the electromagnetic interaction is invariant with respect to time reversal, $\partial_L\partial^*_{L'}$ is real). A numerical table of the F-coefficients, $F_n(LL'J_2J_1)$, is given by Alder, Stech, and Winther.[24] Other notation is

* Since it is assumed that there is no $(S,\ T,\ P)-(V,\ A)$ interference, if all interactions $(S,\ T,\ V,\ A,\ P)$ are considered, the general expression should be written :

$$N\sim|\eta_T|^2\{S\ T\ P\}+|\eta_A|^2\{V,\ A\},$$

where

$$\eta_T=C_T(\textstyle\int i\beta\boldsymbol{\sigma}\times\boldsymbol{r}) \quad\text{and}\quad \eta_A=-C_V(\int\boldsymbol{r}),$$

and $\{STP\}$ denotes the curly bracket of eq. (2·13) evaluated with the coupling constants C_S, C_T, C_P and similarly for $\{VA\}$.

** The factor $2|\eta|^2$ in (2·13) comes from the sum of two contributions we always obtain by making the substitution $C_i\rightleftarrows C_i'$, and by setting $D_j=D_j'=0$. If the D_j's are not zero then we merely add another term to the transition probability identical in form to the r.h.s. of (2·13) with D_i and D_i' substituted for C_i and C_i'.

standard.[25]*

If the γ-ray is not observed ($n=0$), then $\lambda=\lambda'$ (as is clear from the $6j$ symbol, or Racah coefficient of (2·11)) and the expressions simplify:

$$G_{\lambda\lambda}(0)=(-1)^{\lambda}/\sqrt{2\lambda+1}\,, \tag{2·16′}$$

so we can write

$$C=B^{(0)}=\sum_{\lambda}(-1)^{\lambda}b^{(0)}_{\lambda\lambda}/\sqrt{2\lambda+1}\equiv\sum_{\lambda}C_{\lambda}\,, \tag{2·17}$$

$$A^{(0)}=\sum_{\lambda}(-1)^{\lambda}a^{(0)}_{\lambda\lambda}/\sqrt{2\lambda+1}\equiv\sum_{\lambda}A_{\lambda}\,. \tag{2·18}$$

If the circular polarization of the γ ray is measured, the following quantity should be added to the curly bracket $\{\ \}$ of the general distribution function, (2·13):

$$\begin{aligned}(-\tau)\,\{&B^{(1)}(\boldsymbol{p}\cdot\boldsymbol{k})+B^{(3)}(5(\boldsymbol{p}\cdot\boldsymbol{k})^3-3(\boldsymbol{p}\cdot\boldsymbol{k}))/2\\&+A^{(1)}(\boldsymbol{p}\cdot\boldsymbol{s})(\boldsymbol{p}\cdot\boldsymbol{k})+A^{(3)}(\boldsymbol{p}\cdot\boldsymbol{s})(5(\boldsymbol{p}\cdot\boldsymbol{k})^3-3(\boldsymbol{p}\cdot\boldsymbol{k}))/2\\&+E^{(1)}(\boldsymbol{s}\cdot\boldsymbol{p}\times\boldsymbol{k})+E^{(3)}(\boldsymbol{s}\cdot\boldsymbol{p}\times\boldsymbol{k})(3(\boldsymbol{p}\cdot\boldsymbol{k})^2-1)/2\\&+D^{(1)}(\boldsymbol{s}\cdot\boldsymbol{p}\times(\boldsymbol{p}\times\boldsymbol{k}))+D^{(3)}(\boldsymbol{s}\cdot\boldsymbol{p}\times(\boldsymbol{p}\times\boldsymbol{k}))(3(\boldsymbol{p}\cdot\boldsymbol{k})^2-1)/2\}\,,\end{aligned} \tag{2·19}$$

where $\tau=1\,(-1)$ for right (left) circular polarization. The coefficients $B^{(n)}$, $A^{(n)}$, $D^{(n)}$ and $E^{(n)}$ were defined like (2·15). The first two terms have already been measured[11], but the other six require simultaneous measurement of β and γ polarizations, which is very difficult to do at present. We will nevertheless discuss the $E^{(n)}$ terms, since these distributions are odd with respect to the time reversal transformation T, in addition to the $B^{(n)}$ terms. Detailed expressions are not, however, given in the Appendix.

The expressions in the following sections apply to β^- decay. The expressions for β^+ decay are obtained by making the transformation $Z\to(-Z)$, $K_{ik}\to\pm K^{\times}_{ik}$, and $L_{jk}\to\pm L^{\times}_{jk}$, where the upper sign applies if j and k both involve either (S, A, P) or (V, T) interactions, and the lower sign applies if there is interference between (S, A, P) and (V, T) interactions. Thus C_S, C_A, $C_P\to C^*_S$, C^*_A, C^*_P; C_S', C_A', $C_P'\to -C_S'^*$ $-C_A'^*$, $-C_P'^*$; C_V, $C_T\to -C_V^*$, $-C_T^*$; and C_V', $C_T'\to C_V'^*$, $C_T'^*$.

§ 3. The energy spectrum and longitudinal polarization

In this section accurate expressions for the β-energy spectrum and longitudinal polarization will be discussed for each value of $\Delta J=J_0-J_1$. Much of the material to be presented on the spectrum can also be found in older papers such as the classic work of Konopinski and Uhlenbeck[25]. More recent papers which present some of these results are those of Alder, Stech, and Winther[24], Morita and Morita[11] and Curtis and Lewis.[9]

* The notation of this paper is similar to that of Morita and Morita,[11] but is a little different from that of Alder, Stech, and Winther.[24] For example, (2·15) and (2·16) correspond to their equation (6) with our $b_{\lambda\lambda'}{}^{(n)}$ equal to their $+(+4)\dfrac{\sqrt{(2\lambda+1)(2\lambda'+1)}}{|\gamma|^2(1+\gamma_1)}C(\lambda\lambda' n;1\ -1)b_n(\lambda\lambda')$.

Spectrum : Following Konopinski and Uhlenbeck the spectrum is discussed in terms of the correction factor $B^{(0)} \equiv C = \sum C_\lambda$ (see (2·13) — (2·17)). The C_λ's involve only ReK_{ij}, as there is no intereference between parity conserving and non-conserving interactions and similarly for time reversal, i.e there is no interference between terms involving real and imaginary parts of coupling constants.*

Longitudinal Polarization : The longitudinal β-ray polarization, defined by (2·6), is given by

$$P_L = A^{(0)}/C. \tag{3·1}$$

See equations (2·13) — (2·18). Since $(\boldsymbol{p}\cdot\boldsymbol{s})$ is odd under the transformation P (1·2), and even under T (1·3), $A^{(0)}$ involves $Re\, L_{jk}$ and $\alpha Z\, Im\, L_{jk}$.

3A. Matrix elements, including nuclear finite size effects

Exact expressions for all the terms A, B, C, D, E in the β-decay transition probability involve nuclear matrix elements over the electron and neutrino wave functions. In particular, the electron and neutrino radial functions that occur in the matrix element will be of the form

$$g_\kappa(pr)\, j_{l(\nu)}(qr) \quad \text{or} \quad f_\kappa(pr)\, j_{l(\nu)}(qr)$$

where g_κ and f_κ are electron radial functions[23] associated with the angular momentum parameter κ (see discussion following (2·12)), g arising from the large component of the spinor and f from the small component ; and j_l is a spherical Bessel function for the neutrino. The electron radial functions are obtained by solving the Dirac equation for the electron in the field of an extended charge distribution.[27] These functions are not simple. In certain cases essentially numerical expressions have been obtained for the radial functions by various workers.[28-29] In order to exhibit the general properties of the nuclear matrix elements it is not convenient to use such expressions immediately. Instead we observe that

$$g_\kappa(r) \approx (r/\rho)^{l(\kappa)}\, g^P_\kappa(\rho), \tag{3·2}$$

for $r \lesssim \rho$ for conditions of interest in β-decay (and analogously for f_κ). Here $g^P_\kappa(\rho)$ is the point charge radial function evaluated at the nuclear radius ρ. Eq. (3·2) is roughly understood by noting that inside the nucleus the Coulomb force is small. The r^l dependence is then just that expected for small pr. Of course (3·2) is not exact. It is then convenient to write

$$\begin{aligned} g_\kappa(r) &= g^P_\kappa(\rho)\, (r/\rho)^{l(\kappa)} R_g(\kappa, r), \\ f_{-\kappa}(r) &= f^P_{-\kappa}(\rho)\, (r/\rho)^{l(\kappa)} R_f(-\kappa, r), \end{aligned} \tag{3·2a}$$

* Even distributions (0^+) under the transformation T, (1·3), are still expected in general to involve $\alpha Z\, Im\, K_{jk}$ as shown in eq. (1·4). This coulomb correction term appears, however, only when $\kappa' \neq \kappa$ (see (2·11)), because it arises from the phase difference (Δ_κ) between β-ray wavefunctions with different orbital angular momenta (see the definition of Δ_κ in Appendix B). In the case of the spectrum, where $n = \nu_1 = \nu_2 = 0$, $\kappa = \kappa'$ and there is no such interference.

for $r \lesssim \rho$, where, according to (3·2), $R(\kappa)$ is not strongly dependent on r or indeed upon any other variable. Thus, the approximation:

$$R_g = R_f = 1 \tag{3·3a}$$

was used by Fermi, and Konopinski and Uhlenbeck.[5] In order to introduce accurately the effect of the extended nuclear (and screening) charge distribution, it is necessary to consider the detailed character of R_k and R_f.

Let us separate these R's into two parts:

$$\begin{aligned} R_g(\kappa) &= R(\kappa) + \rho R_g^{(1)}(\kappa), \\ R_f(-\kappa) &= R(\kappa) - \rho R_f^{(1)}(-\kappa). \end{aligned} \tag{3·4}$$

The energy independent leading term, $R(\kappa)$, is a rather complicated function of the central potential $V(r)$ for the β-ray and may depend significantly on (r/ρ). Most of its properties are given explicitly by Matumoto and Yamada.[2] If we neglect all higher $(\alpha Z)^2$ terms and the screening effect, we find that

$$\begin{aligned} R(-k) &= 1 - (\alpha Z)^2 O(1/10) + (\alpha Z)^2 (r/\rho)^2 O(1/10), \\ R(k) &= \frac{3k}{2k+1} - \frac{k}{2k+3}\left(\frac{r}{\rho}\right)^2 + O(\alpha Z)^2, \end{aligned} \tag{3·3b}$$

for the uniform charge distribution, $V(r) = -(\alpha Z/2\rho)(3-(r/\rho)^2)$ for $r<\rho$. Here $k=|\kappa|$.*

The next terms, $R_g^{(1)}(\kappa)$ and $R_f^{(1)}(\kappa)$, are less than $1/\alpha Z$ for $\kappa>0$ and of order αZ for $\kappa<0$. Both $R^{(1)}$'s are energy dependent. The corrections due to the $R^{(1)}$'s are generally small, i.e. about 1%, but they become important in special cases such as RaE. We will discuss these contributions in Appendix B.

The radial functions are now written:

$$g_\kappa(r)\, j_l(qr) = g_\kappa^{\rho}(\rho)\,(r/\rho)^{l(\kappa)}\, j_l(qr)\, R(\kappa, r). \tag{3·2b}$$

It is not convenient to use such expressions in their entirety. Instead, since $q\rho$, $p\rho \ll 1$, where ρ is the nuclear radius, it is desirable to expand the lepton functions in terms of these quantities (actually $q^{l+1} F_\nu(Z, W)$ is expanded). Cutting off such an

* In Matumoto and Yamada's notation,

$$R(-k) = \nu\sqrt{2k_l/(k+\gamma_k)}\, I_k(r), \tag{3·4a}$$

$$R(k) = \nu\sqrt{2(k+\gamma_k)/k}\,[-(\rho/\alpha Z)\, C_k(r)], \tag{3·4b}$$

where

$$C_k(r) = kr^{-2k-1}\int_0^r dx\, x^{2k} V(x)\, I_k(x). \tag{3·4c}$$

See eq. (13) and Figs. 3 to 6 of their paper.[2] They also show results for the surface charge distribution, $V(r) = -\alpha Z/\rho$ for $r<\rho$. We find in the same approximation:

$$\begin{aligned} R(-k) &\sim 1 + O((\alpha Z)^2), \\ R(k) &\sim 2k/(2k+1) + O((\alpha Z)^2). \end{aligned} \tag{3·3c}$$

expansion in ρ for the transition probability at the zeroth power of ρ should be an excellent approximation. Thus we assume:

Assumption II: *All terms in the final expressions of A, B, C, D, and E involving positive powers of ρ, can be neglected. The error involved will be, at most, of relative size ρ, or order of magnitude 1%.*

This assumption will be fully defined in terms of the notation we will use. It will become clear that it means that one keeps at least one term in which the coefficient of ρ^0 is not abnormally small. The expressions are still complicated in this approximation but the energy dependence is now explicitly exhibited.

The nuclear matrix elements now have the form

$$(\Psi_f, O_\lambda r^n R(\kappa) \Psi_i) \equiv \int O_\lambda r^n R(\kappa),$$

where there are six spin and angle operators O_λ, of rank λ, that can contribute to first forbidden decays. (There are other smaller operators that contribute, which are called "third forbidden". They are discussed in Appendix A). In any decay the ranks λ that may contribute obey the relation:

$$|J_0 - J_1| \leq \lambda \leq J_0 + J_1. \tag{3·5}$$

The six general types of matrix elements associated with the O_λ are shown, and named in Table 2. The parameters $u_n(\kappa), \cdots, z_n(\kappa)$ which we introduce will be the ratios of

Table 2.

λ	Operator	Nuclear Matrix Element and Parameter
0	$w = (\beta)\boldsymbol{\sigma}\cdot\boldsymbol{r}$	$\eta w_n(\kappa) = C_{A(T)} \int w(r/\rho)^n R(\kappa)$
0	$v = i(\beta)\gamma_5$	$\eta \xi' v_n(\kappa) = C_{A(P)} \int v(r/\rho)^n R(\kappa)$
1	$v = i(\beta)\boldsymbol{\sigma}\times\boldsymbol{r}$	$\eta u_n(\kappa) = C_{A(T)} \int u(r/\rho)^n R(\kappa)$
1	$x = (\beta)\boldsymbol{r}$	$\eta x_n(\kappa) = -C_{V(S)} \int x(r/\rho)^n R(\kappa)$
1	$y = i(\beta)\boldsymbol{\alpha}$	$\eta \xi' y_n(\kappa) = -C_{V(T)} \int y(r/\rho)^n R(\kappa)$
2	$z = (\beta) B_{ij}$	$\eta z_n(\kappa) = C_{A(T)} \int z(r/\rho)^n R(\kappa)$

Definition of "nuclear parameters" u, v, w, x, y, z*; where $\xi' \equiv 1/4\rho$ (see (3·8)). If the interactions are invarient under time reversal all the parameters are real. In this paper ρ is equal to the nuclear radius. It is not a parameter. Third forbidden matrix elements that may contribute are discussed in Appendix A.

* Our definitions of the γ-matrix (see footnote following (2·1)) and of the lepton wavefunction are such that the matrix element $(\int\boldsymbol{\alpha})$ has the opposite sign from that in Konopinski and Uhlenbeck[25] and most other papers. The matrix elements involving the D_i's are of the same form and occur in the same way in the transition probability as the above, as discussed in the footnote following eq. (2·16). The signs of $x_n(\kappa)$ and $y_n(\kappa)$ are chosen to conform with other papers[11].

the various matrix elements compared to a standard matrix element γ_i, so that γ_i^2 can be taken out as a common factor in the transition probability (as in (2·13)).

By including the R functions in the nuclear parameters we have a formulation general enough to include finite size effects accurately. It is not necessary to specify the form of the R's exactly, because they are associated under the integral sign with unknown nuclear radial functions. As a result, for each operator, e.g. x, there may occur different parameters for each κ. This situation can be simplified. It is quite a good approximation from (3·3b) to set

$$R(\kappa)=R(-1) \quad \text{for all } \kappa, \text{ except } \kappa=1.$$

Then for example only the parameteos $x(1)$ and $x(-1)$ are independents. An example is discussed in Appendix C.

If time reversal is not violated in strong interactions any products of two (reduced) nuclear matrix elements are real (in this case we can without loss of generality consider the nuclear matrix elements to be real). We shall use the following assumption :*

Assumption III. *The strong interaction is invariant under time reversal.*

Violation of time reversal invariance in the β-decay will, however, cause some parameters of Table 2 to be complex, due to complex coupling constants. Thus, for the (S, T, P) combination it will be convenient to set :

$$u_0(-1)=1, \quad \eta=C_T\int i\beta\boldsymbol{\sigma}\times\boldsymbol{r}R(-1), \tag{3·6}$$

then x and v may be complex while w, y, z are real. For the (V, A) combination (especially for $\lambda=1$) it is convenient to set :

$$x_0(-1)=1, \quad \eta=-C_V\int \boldsymbol{r}R(-1), \tag{3·7}$$

then y is real and u may be complex. We expect the absolute value of these nuclear parameters to be of order of magnitude unity (except for P interaction).

3B. The ξ-expansion

In Appendix C the quantities C_λ and A_λ needed to predict C and P_λ are given to an accuracy consistent with all assumptions discussed above. (We do in fact omit some terms which should technically be included under Assumption II. These omissions are discussed in Appendix A.) In order to exhibit these expressions in a convenient form, we have followed the custom[31] of introducing the parameter ξ instead of ρ ; and we have arranged terms in order of decreasing powers of ξ, where*

$$\xi=\alpha Z/2\rho\approx Z/A^{1/3}, \quad \xi'=1/4\rho. \tag{3·8}$$

* In order that the nuclear parameters $y(\kappa)$ and $v(\kappa)$ be of order unity we introduce the ξ' parameter. We have assigned ξ' the nominal value of 1.4ρ. We do not actually know how to determine ξ' reliably, although there have been some attempts to do so,[9] but we feel that there is some empirical evidence for setting $\xi'\sim\xi$ for heavy nuclei and that ξ' should not become small for light nuclei.[30]

We shall use $\tilde{\xi}$ to designate either $\tilde{\xi}$ or $\tilde{\xi}'$ in our general discussion. Aside from very high energy decays, $\tilde{\xi}$ is the only quantity in the expressions which is expected to be an order of magnitude greater than unity. Thus we anticipate for most decays that

$$\tilde{\xi} \gg W_0 \quad \text{in addition to} \quad \tilde{\xi} \gg 1 ,$$

and that the order of magnitude of the nuclear parameters u, v, w, x, y, $z(\kappa)$ is unity. As a result we write the shape factors C_0 (and C_1) in the form:

$$C_0 = C_0(\tilde{\xi}^2) + C_0(\tilde{\xi}) + C_0(1) , \tag{3·9}$$

and similarly for other quantities. The series is cut off at the term of order $\tilde{\xi}^0 = 1$ in accordance with Assumption II.

The $\tilde{\xi}$-approximation: This series provides a basis for a convenient approximation. We will call the approximation in which only the leading term in powers of $\tilde{\xi}$ in any expression is kept, the "*$\tilde{\xi}$-approximation*". The accuracy of this approximation is discussed below.

Arranging terms according to the powers of $\tilde{\xi}$ we find that positive powers of $\tilde{\xi}$ occur in connection with certain nuclear matrix element combinations. Thus for $\lambda = 0$, $\tilde{\xi}$ only occurs in the combination:

$$V = \tilde{\xi}' v_0(-1) + (2/1+\gamma_1)\tilde{\xi} w_0(1), \tag{3·10}$$

and for $\lambda = 1$ in the combination:

$$Y = \tilde{\xi}' y_0(-1) - (2/1+\gamma_1)\tilde{\xi}(u_0(1) + x_0(1)). \tag{3·11}$$

Thus the shape factors have the form

$$\begin{aligned} C_0 &= |V|^2 + O(V) + O(1), \\ C_1 &= |Y|^2 + O(Y) + O(1), \\ C_2 &= \qquad\qquad O(1). \end{aligned} \tag{3·12}$$

The detailed expressions are given in Appendix C.

In the case of the general non-unique first forbidden decay, there may be contributions from all ranks $\lambda = 0$, 1, 2 of nuclear matrix elements depending on the nuclear spin J_0 and J_1 (3·5). Therefore, we should consider all C_λ's (3·12), in the correction factor for the energy spectrum. The $\tilde{\xi}^2$ terms are energy independent, while the $\tilde{\xi}$ and $\tilde{\xi}^0$ terms are energy dependent. In the $\tilde{\xi}$-approximation only the $|V|^2$ and $|Y|^2$ terms are present, and the correction factor is energy independent (i.e. the shape is the allowed transition shape).

Strictly speaking, we see that the correction factor is not an expansion in $\tilde{\xi}$ but essentially in Y, for example. It is possible, though very unlikely because of the size of $\tilde{\xi}$, that $|Y| = O(1)$. If such a strong cancellation (among the nuclear matrix elements in Y) occurs, we cannot use the $\tilde{\xi}$-approximation at all but have to use the entire expression

represented by (3·12). The shape factor for such a case of strong cancellation will generally be quite differnt from a constant. The best example of this, RaE, is discussed below. Energy dependence is also expressed, if one nuclear parameter is relatively large, e.g. $z_0(\kappa) \approx Y \gg x_0(\kappa)$ by some selection rule effect. These cases are discussed in a forthcoming article.

Leading aside such cases which may be unusual, we can say that the $\tilde{\xi}$ approximation is a good first approximation for the nonunique decays. The corrcetions to the $\tilde{\xi}$-approximation will vary widely in magnitude being of order 1% to order 10%. Thus spectrum shape experiments of the highest accuracy now possible should usually reveal some deviation from an "*allowed shape*". Of course the measurement of these corrections gives us new information about the decay.

In first approximation we then have the shape factor energy dependence:

$$C \sim 1, \qquad \text{(i.e. constant)} \quad (3\cdot13)$$

corresponding to the $|V_1|^2$ and $|Y|^2$ terms in (3·12) ($C_\lambda(\tilde{\xi}^2)$ term). In the next approximation, including terms of order $\tilde{\xi}^2$ and $\tilde{\xi}$ but not of order 1 (*i,e,* $C_\lambda(\tilde{\xi}^2)$ and $C_\lambda(\tilde{\xi})$), we have a correction factor of form (see Appendix C, eqs. (C–6, 7, 14, 15)):

$$C \sim (1 + aW + b/W). \qquad (3\cdot14)$$

Treated as a two parameter expression, (3·14) should provide an excellent fit with

$$a,\ b \ll 1, \qquad (3\cdot15)$$

except extreme cases such as RaE. The test of validity of (3·14) is (3·15). If a fit to a particular experiment allows a whole range of parameters a and b the *a priori* best fit corresponds to assuming a and b small. If (3·14) with (3·15) is not adequate then the full expression for C, represented by (3·12), should be used. Then the shape factor should be of the form (see Appendix C, eqs. (C–6, 7, 8, 14, 15, 16, 21)):

$$C \sim 1 + aW + b/W + cW^2. \qquad (3\cdot16)$$

Because of the mathematical flexibility of the complete theoretical predictions represented by (3·12) or (3·16), we advocate the philosophy of trying to analyze the experimental data with the assumption that no unusual cancellation among, or huge contributions from, nuclear matrix elements occur. Thus, if of allowed form, the shape should be interpreted in terms of the $|V|^2$ and $|Y|^2$ terms in C (i.e. the $\tilde{\xi}$-approximation). The experimental shape, if known to an accuracy greater than 1 $\tilde{\xi}$ or if not of allowed form, should be empirically fit in terms of the simple expression (3·14). (3·16), with its three parameters, should be used in case of necessity. Determination of the nuclear matrix elements by comparison of the empirical a and b of (3·14) with the theoretical expression involving $C_\lambda(\tilde{\xi}^2)$ and $C_\lambda(\tilde{\xi})$ terms, (but not terms of order 1) can then be carried out when convenient. In particular we expect that useful determinations of nuclear parameters will in future be obtained by combining results of several types of experiments involving in the same decay.

In the $\tilde{\xi}$-approximation the longitudinal polarization for the non-unique decay is:

$$P_L = -(p/W). \tag{3·17}$$

When the next terms in the $\tilde{\xi}$-expansion are included, P_L is of form:

$$P_L = -(p/W)[1+(1/W)(R_1/C)+\alpha Z(1/p)(I_1/C)]. \tag{3·18}$$

The R_1 and I_1 arise from the $A_\lambda(\tilde{\xi})$ as given in Appendix C, eqs. (C–7, 15). The constant terms R_1 and I_1 include contributions from ReL_{jk} and ImL_{jk}, respectively, and the form of C is given by (3·14).

3C. Special cases in non-unique transitions

The case of the $1-0$ *transition*: The famous example in this class is the RaE decay. Here $\lambda=1$ only. The shape correction factor is $C=C_1$. The shape deviates very strongly from the allowed type.[33] This is explained by assuming $|Y|^2 \ll \tilde{\xi}^2$, i.e., there is a strong cancellation of matrix elements such that $|(\tilde{\xi}'/\tilde{\xi})y_0(-1)-(2/1+\gamma_1)\cdot(u_0(1)+x_0(1))| \ll 1$. The $\tilde{\xi}$-approximation cannot be used even as a first approximation and several interesting features appear in the more complicated expression for C_1, (3·16). The correction factor in all its detail is just the sum $C_1(\tilde{\xi}^2)+C_1(\tilde{\xi})+C_1(1)$ as given in Appendix C, eqs. (C–14), (C–15), and (C–16). It is seen, for example, that there are terms involving the first power of q and, as shown in § 1, these terms have opposite sign for (S, T, P) and (V, A) combinations. There is also some possibility to determine the ratio $ReK_{ij}/|K_{ij}|$.*

Quantitative examination of the shape factor requires careful examination of finite size effects, as discussed above, so that, for example $y_0(1)$ and $y_0(-1)$, or $u_0(1)$ and $u_0(-1)$ are considered as more or less independent parameters (depending on one's opinions about nuclear radial functions). We shall discuss this in Appendix B and C. In addition we must consider the high angular momentum expansion terms in the lepton wave functions known generally as third forbidden contributions. We shall discuss this problem in Appendix A. These contributions are included in the expressions of Appendix C.

There are many unknown nuclear parameters. (If time reversal does not hold there is one additional parameter, essentially the phase difference between C_A and C_V or between C_S and C_T.) We need a very accurate spectrum measurement, or other experiments, to determine these parameters. Such information may be obtained by measuring the longitudinal polarization in the RaE decay. The longitudinal polarization is

$$P = A_1/C_1, \tag{3·19}$$

where A_1 has the same structure (3·12) as C_1. The leading term of A_1 is $(-p/W)|Y|^2$, so that in the $\tilde{\xi}$-approximation

---

* It has already been shown by Lewis[34] and by Fujita et al.[34] that ImY cannot be large, on the basis of the spectrum shape. The essential point of their discussion is as follows: consider the (V, A) combination, then we have the parameters $Reu_0(1)$, ReY, and ImY $(=-(2\xi/(1+\gamma_1))\ Im\ u_0(1))$. The leading term in C_1 is $|Y|^2=(ReY)^2+(ImY)^2$. In order that the shape is not allowed, $|Y|^2=0(1)$. Thus $Im\ u_0(1)$ must be small, i.e. $(\xi\ Im\ u_0(1))=0(1)$ (actually $Im\ u_0(1)/Re\ u_0(1)=0(1/\xi)$). A corresponding agrument can be made in the (S, T, P) case.

$$P_L = -p/W. \tag{3·17}$$

The energy dependence of the other terms in A_1 is quite different from the leading term and from each other. There is some indication that (3·19) does not hold.[37] This is clearly to be expected since the $\tilde{\xi}$-approximation didn't apply to the spectrum and both the spectrum and the polarization term A have the same structure. It is of interest to exhibit the form of eq. (3·18),

$$P_L = -(p/W)\left\{1+\frac{1}{C}\left[\frac{R_1}{W}+\alpha Z\frac{I_1}{p}+R_2+\alpha Z I_2\frac{W}{p}\right]\right\}, \tag{3·20}$$

where R and I involve the real and imaginary parts of L_{ij}, respectively. They are energy independent and depend on nuclear parameters. The actual expression represented by the R's and I's are found in Appendix C, eqs. (C-14), (C-15), and (C-16). In particular R_1 has the simple form,

$$R_1/C = -b/(1+aW+b_0W+cW^2), \tag{3-21a}$$

where a, b, c were introduced in (3·16). The spectrum shape measurements of Plassman and Langer[33] in principle allow us to determine a, b, and c. However, since we do not have a very accurate knowledge of the maximum energy, W_0, we have to treat W_0 as an adjustable parameter. Thus, we have 4 unknown parameters to describe the shape alone, and, at present, more than one set of parameters is satisfactory, say sets with $b>1$ and $b<1$. The value of R_1 will help to decide the experimental valves of a, b, and c. This is generally true for every β decays. There is another point of immediate interest. For small p especially, we see an opportunity to test time reversal. The polarization involving the imaginary coupling constant does not go to zero with p. We may expect that I_1 and I_2 are measurable. We hope that an experimental attempt to detect this term will be made.

The case of the 0-0 transition (with parity change): We discuss this case principally to see whether it is possible to detect the pseudoscalar interaction. It is well known that the nuclear matrix element $(\int_1\vec{\gamma}\gamma_5)$ is expected to be very small and that it should depend on the lepton momenta.[9] In spite of its small size it is of interest to look experimentally for the quantity $r_0(\kappa)\sim C_{T'}\int_1\vec{\gamma}\gamma_5 R(\kappa)$: (1) Consider longitudinal polarization in any $\Delta J=0$ transition. If the spectrum has an allowed shape, then a longitudinal polarization $P_L=-(p/W)$ for electrons would indicate that v is very small, or that

$$C_p = -C_p'. \tag{3·21}$$

(2) Consider the spectrum shape in the 0-0 transition. Since the leading term in C_0 (i.e. $C_0(\tilde{\xi}^2)$) is

$$C_0 = |V|^2 = |\tilde{\xi}'\nu_0(-1)+\tilde{\xi}w_0(1)\,(2/1+\gamma_1)|^2, \tag{3·22}$$

in the STP case, simply the observation of substantial deviation from allowed shape would mean that r and w are comparable. This would detect C_p. In the (V, A, P) case, since there is no $(V, A)-P$ interference, if (3·21) holds, the leading term in C_0 has the form

$$C_0 = |V_1|^2+|V_{V-A}|^2. \tag{3·23}$$

Observation of substantial deviation from allowed shape may again indicate a contribution from the *P*-interaction. The analysis becomes very difficult, however, because energy dependence can also arise from the next term in the $\hat{\xi}$ expansion.

The 0-0 transition may also be employed to check whether the (V, A) combination is more favorable than the (S, T, P) combination if we can assume that the *P*-interaction can be neglected (or conclude experimentally, that the energy spectrum in this transition has the allowed shape precisely). In order to simplify the argument, we assume time reversal invariance and $(\alpha Z)^2 \ll 1$. The correction factor can be written in the form of (3·14):

$$C_0 = |V|^2 k (1 + aW + b/W)$$

where

$$ak = (2w_0(1)/3V)[1 \pm (w_0(-1)/w_0(+1))], \tag{3·24}$$

$$bk = -2w_0(1)/3V, \tag{3·25}$$

$$k = 1 \mp W_0(2w_0(-1)/3V) \tag{3·26}$$

(see Appendix C, (C–6) and (C–7)). The upper and lower signs refer to the (S, T, P) and (V, A) combinations, respectively. In the (S, T, P) case, since it is assumed that the *P* interaction $v(\kappa)$, does not contribute appreciably, then a and b depend only on the matrix elements $w(\kappa)$ associated with the tensor interaction. Thus, for example, neglecting the v term above

$$a = (2/3\hat{\xi})[1 + w']/[1 - (2W_0 w'/3\hat{\xi})]. \tag{3·27}$$

Here $w' = w_0(-1)/w_0(1)$ should be a positive quantity larger than unity, (3·3b) (See Appendix C). Therefore a is fairly well fixed to be approximately $1/\hat{\xi}$. This is an observable effect. In the case of the (V, A) combination we have two comparable independent matrix elemnts, $w(\kappa)$ and $v(\kappa)$, so that a and b may assume a wide range of values.

3 D. The unique transition

In the case of the so-called unique first forbidden transition where $\Delta J = 2$ with parity change, only the matrix element of rank $\lambda = 2$ contributes. The main contribution is given by B_{ij}, i.e. $z_0(\kappa)$ listed in Table 2, an order of unity term in the $\hat{\xi}$-expansion. Then, the correction factor $C = C_2(1)$ is:

$$C_2 = (1/12)[q^2|z_0(-1)|^2 + p^2\lambda_1|z_0(-2)|^2], \tag{3·28}$$

where*

$$\lambda_1 = (2 + \gamma_2) F_1(W, Z)/[2(1 + \gamma_1) F_0(W, Z)], \tag{3·29}$$

---

* Numerical values of λ_1 can be found from the functions L_0 and L_1 tabulated in Siegbahn,[37] Appendix III p. 884, since,

$$\lambda_1 \simeq 9L_1/p^2 L_0.$$

See also Fig. 1 of Davidson,[37] where $\lambda_1 = A$.

Here γ_k and the Fermi factor, $F_k(W, Z)$, are defined by (2·14) and in Appendix B, (B·25). As we discussed in (3·3b) the ratio $(z_0(-2)-z_0(-1))/z_0(-1))$ is order of 1%. There are some correction terms of order 1 10, which are the interference terms between the first and the so-called third forbidden transition. However, the same type of energy spectrum is expected. We can just change the definition of $z_k(k)$. This will be shown in Appendix A, (A·10), and it is found that the $q^2+\lambda_1 p^2$ spectrum shape is retained (to accuracy $\alpha Z_1 \rho$ or about 1%, if B_{ij} is not abnormally small).

The longitudinal polarization is

$$P_L = A_2 \ C_2 = \mp p \ W \tag{3·30}$$

for $\beta^{\mp}$ particles, respectively.

§ 4. The β-γ correlations

In this section we will give results for six measurable quantities involving β-γ correlation, with and without measuring the β and γ polarizations. These quantities discussed here are reviewed in Table 3 (next section). Some of this material has already been presented by Alder, Stech and Winther,[3] by Curtis and Lewis[5] and by Morita and Morita.[11]

It is of interest to note that in a general β-γ correlation experiment the energy independent quantity $G_{\lambda\lambda'}(n)$ (see eq. (2·16)) where λ may not equal λ', will occur. Thus an accurate experiment can give information on the competition between γ-rays of different multipole order, just as can be found in γ-γ correlations. If the γ-ray is a pure multipole then $G_{\lambda\lambda'}$ is just a numerical constant depending on the various angular momenta involved and independent of nuclear matrix elements.

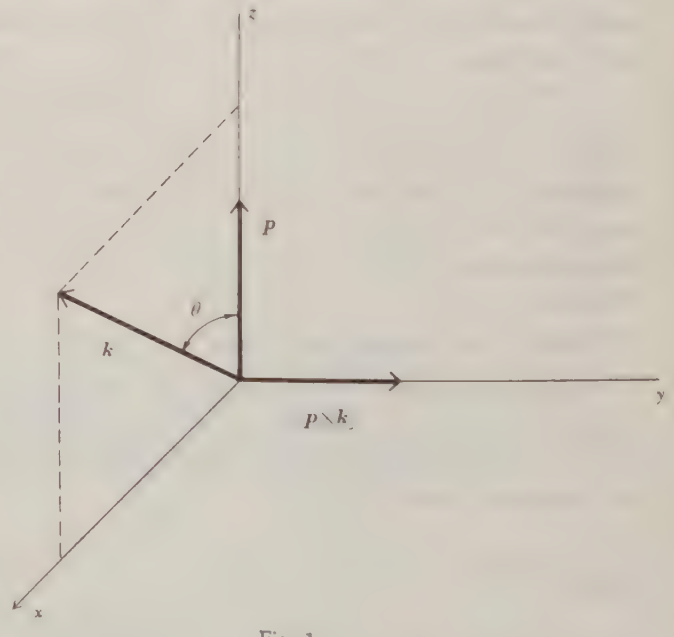


Fig. 1.

Let us first consider the simpler experiments of β-γ correlation with and without γ-circular polarization or longitudinal β polarization. In considering these distributions let us speak in terms of a definite coordinate system in which $\boldsymbol{p}$ lies along the z-axis, $\boldsymbol{k}$ in the xz plane and $[\boldsymbol{p}\times\boldsymbol{k}]$ along the y axis as shown in Fig. 1.

β-γ directional correlation: If neither the polarizations of β or γ are measured the distribution will have the form:

$$N=1+\varepsilon\left(\tfrac{3}{2}\cos^2\theta-\tfrac{1}{2}\right), \tag{4·1}$$

where θ is the angle between β-ray and γ-ray, and

$$\varepsilon=B^{(2)}/C, \tag{4·2}$$

(see (2·13)). Being even under the transformations T and P ((1·2) and (1·3)), this distribution involves ReK_{kj} and $\alpha\mathcal{Z}ImK_{kj}$.

β-γ correlation measuring circular polarization: In this case we define the γ-ray polarization analogously to the longitudinal β-polarization (2·6):

$$P(\gamma)=\frac{N(R)-N(L)}{N(R)+N(L)}=\omega\left(\frac{p}{W}\right)\cos\theta. \tag{4·3}$$

The polarization as a function of θ is then:

$$\omega=\left(-\frac{W}{p}\right)\frac{B^{(1)}+B^{(3)}\left(\frac{5}{2}\cos^2\theta-\frac{3}{2}\right)}{C+B^{(2)}\left(\frac{3}{2}\cos^2\theta-\frac{1}{2}\right)}. \tag{4·4}$$

The polarization will involve ReL_{kj} and $\alpha\mathcal{Z}ImL_{kj}$ in the numerator.

β-γ correlation measuring longitudinal polarization: The longitudinal polarization as a function of θ is:

$$P_L^{\gamma}=\frac{A^{(0)}+A^{(2)}\left(\frac{3}{2}\cos^2\theta-\frac{1}{2}\right)}{C+B^{(2)}\left(\frac{3}{2}\cos^2\theta-\frac{1}{2}\right)}. \tag{4·5}$$

The unique first forbidden transition: The unique first forbidden decay has quite different properties from other first forbidden decays. Its particular distinction is that the various quantities are essentially independent of nuclear β-decay matrix elements, (as we shall discuss in Appendix A, if we include the so-called third forbidden transition, there is some minor modification for the nuclear matrix element). As it is not important for the energy dependence of the following quantities, we assume that $z_0(-1)=z_0(-2)$. (See the discussion in § 3·D, especially (3·3b).). The β-γ correlation is

$$\varepsilon=\frac{1}{C_2}b_{22}{}^{(2)}G_{22}(2)=-\sqrt{\frac{7}{2}}G_{22}(2\,;\,J_0J_1J_2)\frac{p^2\lambda_1}{q^2+p^2\lambda_1}, \tag{4·6}$$

where λ_1 is defined in eq. (3·29). If there is competition between γ-rays of different multipole orders the matrix elements for the γ-emission occur in G_{22}. The circular polarization as a function of the angle between β and γ is also independent of the β-decay matrix element,

$$\omega=\frac{(5q^2+3\lambda_1p^2)G_{22}(1)-6\lambda_1p^2\left(\frac{5}{2}\cos^2\theta-\frac{3}{2}\right)G_{22}(3)}{\sqrt{10}(q^2+\lambda_1p^2)-\sqrt{35}\lambda_1p^2\left(\frac{3}{2}\cos^2\theta-\frac{1}{2}\right)G_{22}(2)}. \tag{4·7}$$

The longitudinal polarization as a function of angle between β and γ is independent of the angle:

$$P_L=-p/W. \tag{4·8}$$

As we showed, we have no unknown nuclear parameter in a good approximation.

The absence of nuclear β-decay matrix elements means that the unique transition could be employed to calibrate experimental shape determinations for the above processes. The above expressions are sufficiently accurate for such purposes, i.e. they should be accurate to about 1% for small W_0.

The non-unique first forbidden transitions: In order to discuss the various experiments with some ease, we will present results in this section in the $\tilde{\xi}$-approximation, keeping only the leading terms in powers of $\tilde{\xi}$. Considering the difficulties of measurements in these cases, this approximation should usually be satisfactory (i.e. if there is no large cancellation). In general, this correction is of order $(1/\tilde{\xi}^2)=4/\rho\approx 3\alpha Z/A^{1/3}$. Even for the heavy nuclei, say $A^{1/3}\approx 6$, this correction term is 10%. The reader is referred back to the discussion in § 3 about the application of the $\tilde{\xi}$-approximation to the non-unique decay (see (3·9) etc.). Corrections to the $\tilde{\xi}$-approximation will be discussed in general terms here, but will be given explicitly in Appendix C. We neglect all next correction terms (i.e. those of relative order $1/\tilde{\xi}^2$ or less). Thus, as illustrated in Table 3, we neglect correction terms which are of order $(1/\tilde{\xi})^2$ for the circular polarization, ω, and the longitudinal polarization P_L, and of order $(\alpha Z)^2$ for all other distributions.

The β-γ correlation in the $\tilde{\xi}$-approximation has the form:

$$\mathcal{E}=\lambda_2(p^2/W)\,(R_3/C)+\alpha Z\lambda_3 p(I_3/C), \tag{4·9}$$

where the λ_i's are defined in Appendix C, (C·29) to (C·30). In the approximation where $(\alpha Z)^2\ll 1$, $\lambda_i\approx 1+O((\alpha Z)^2)\,(W/p)^2$. The R_3's and I_3's are energy independent and include ReK_{ij} and ImK_{ij}, respectively. They are given in terms of $b_{ij}^{(k)}(\tilde{\xi})$ in Appendix C, (C·22, 24, 26). In particular, neglecting the interference between the first and the third forbidden transitions, which will be discussed in Appendix A,

$$\begin{aligned}R_3=&\sqrt{2/3}\cdot\{G_{02}(2)\,Re[z_0(-2)V^*]\\&-2G_{11}(2)\,Re[(x_0(-2)-\tfrac{1}{2}u_0(-2))\,Y^*]\}-G_{12}(2)\,Re[z_0(-2)\,Y^*],\end{aligned} \tag{4·10}$$

$$\begin{aligned}I_3=&\sqrt{3/2}\cdot\{\tfrac{1}{2}G_{02}(2)\,Im[z_0(-2)V^*]\\&-G_{11}(2)\,Im[(x_0(-2)-\tfrac{1}{2}u_0(-2))\,Y^*]\}-\tfrac{3}{4}G_{12}(2)\,Im[z_0(-2)\,Y^*].\end{aligned} \tag{4·11}$$

C is the energy correction factor. Since C is of order $\tilde{\xi}^2$, (see (3·9), for example),

$$\mathcal{E}=O(1/\tilde{\xi}) \tag{4·12}$$

As the experiment is relatively easy, it provides perhaps the best possibility for determining the imaginary parts of the coupling constants among the experiments discussed here. In order to do this experiment, we should check the reliability of the $\tilde{\xi}$-approximation for the particular decay. Among other ways, we can estimate the correction to the leading term (4·9) in the $\tilde{\xi}$-expansion by measuring the coefficient, $\mathcal{E}$, for high energy β-rays, where $p\approx W$. The energy dependence of λ_i can be neglected at high energies because they depend on W/p; so, in the $\tilde{\xi}$-approximation, $\mathcal{E}$ has the energy dependence:

$$\mathcal{E}\sim W \quad \text{for high energy } \beta\text{-rays.} \tag{4·13}$$

On the other hand, the correction terms to (4·9) have the form

$$[(p^2/W)(R_3'/C)+p^2(R_3''/C) \qquad (4\cdot14)$$
$$+\alpha Zp(I_3'/C)+\alpha Z(p/W)(I_3''/C)$$
$$+\alpha Zpw(I_3''/C)],$$

where R_3', I_3' and so on, are energy independent to order $(\alpha Z)^2$, and are of order unity. They can be obtained from $b^{(2)}_{\lambda\lambda'}(1)$ in Appendix C, eqs. (C·23 to 28). The correction factor C must also be given to an accuracy consistent with the accuracy of $\mathcal{E}$. Thus to include $(1/\tilde{\xi})$ corrections in $\mathcal{E}$ we must use (at least) $\tilde{\xi}^2$ and $\tilde{\xi}$ order terms for C in (4·9) (such as represented by $(1+aW+b/W)$, (3·14).

The β-circularly polarized γ correlation is:

$$\omega=\frac{1}{NC}(2G_{01}(1)\,ReVY^*-2^{1/2}G_{11}(1)\,|Y|^2), \qquad (4\cdot15)$$

N is defined in (4·1). Since $NC=C=0(\tilde{\xi}^2)$, in the $\tilde{\xi}$-approximation,

$$\omega=O(1). \qquad (4\cdot16)$$

For $\Delta J=1$ the result depends only on nuclear spins and possible γ-ray mixing:

$$\omega=-\sqrt{2}\,G_{11}(1). \qquad (4\cdot17)$$

Meanwhile for $\Delta J=0$ and a pure multipole γ-ray (or known γ-ray mixing) ω can be used to determine the sign of $ReVY^*$. Terms in imaginary parts of the coupling constants can enter only to order $(1/\tilde{\xi})$ smaller than (4·15). Angular dependence in the numerator (i.e. $B^{(3)}$), enters to order $(1/\tilde{\xi}^2)$ smaller than (4·15). Angular variation of ω will occur, however, in association with the β-γ directional correction coefficient $\mathcal{E}$ in the denominator. The $(-p/W)$ circular polarization energy dependence provides yet another check of the $\tilde{\xi}$-approximation for a particular decay. In particular, the $(1/\tilde{\xi})$ corrections to ω are of the form:

$$(1/CN)(R_4+WR_4'+\alpha Z(1/p)I_4+\alpha ZpI_4'), \qquad (4\cdot18)$$

where (R_4/NC) and (I_4/NC) are of order $(1/\tilde{\xi})$ and are energy independent to order $(\alpha Z)^2$. They involve real and imaginary parts of the coupling constant combinations, respectively. The detailed expression represented by and the R's and I's is given in terms of $b^{(1)}_{\lambda\lambda'}(\tilde{\xi})$ in Appendix C, eqs. (C·33 to 40). Again we note that to include the $(1/\tilde{\xi})$ correction in ω requires use of at least $\tilde{\xi}^2$ and $\tilde{\xi}$ order terms in NC in (4·15).

The longitudinally polarized β-γ correlation in the $\tilde{\xi}$ approximation is simple:

$$P_L=-p/W, \qquad (4\cdot19)$$

independent of the matrix elements. Angular dependence will enter to order $1/\tilde{\xi}$ smaller than this, as will different energy dependence. The angular independent longitudinal polarization was discussed in some detail in § 3C and 3D. The inclusion of the $(1/\tilde{\xi})$ correction to (4·19) gives an expression whose energy dependence has the form:

$$P_L=(-p/W)\{1+(1/CN)[(1/W)R_1+\alpha Z(1/p)I_1]$$
$$+(1/CN)[(1/W)R_5+\alpha Z(1/p)I_5+pI_5'](\tfrac{3}{2}\cos^2\theta-\tfrac{1}{2})\}, \qquad (4\cdot20)$$

where (R/NC) and (I/NC) are of order $(1/\tilde{\xi})$. The detailed expression represented by the R's and I's is given in terms of $A_\lambda(\tilde{\xi})$ and $a^{(\sigma)}_{\lambda\lambda'}(\tilde{\xi})$ in Appendix C, (C·2), (C·5), and (C·42 to 45).

Now let us consider transverse β polarization measurements. The transverse polarization is associated with terms of type D and E of Table 1. We have some interest in these terms because they may provide a good means of testing time reversal. There are two types of transverse polarization measurements: in the plane of β and γ and perpendicular to that plane. In addition in each of these cases we may measure circular polarization of the γ, if we wish.

Transverse β polarization in the plane of β and γ: In the plane, the polarization (with no measurement of γ polarization) will be (using a definition analogous to the longitudinal polarization, (2·6) and using Fig. 1)

$$P_{T\parallel}=\frac{D^{(2)}\sin\theta\cos\theta}{C+B^{(2)}(\frac{3}{2}\cos^2\theta-\frac{1}{2})}. \qquad (4\cdot21)$$

This is an even distribution (0) under the transformation P and T. In the non-unique first forbidden transition, $P_{T\varepsilon}$ has the form:

$$P_{T\parallel}=\sin\theta\cos\theta\left\{-\tfrac{3}{2}\lambda_6\frac{p}{W}\frac{R_3}{CN}-2\alpha Z\lambda_7\frac{I_3}{CN}\right\}. \qquad (4\cdot22)$$

This expression strongly resembles that for the simple directional correlation, ε. The R_3 and I_3 are exactly the same quantities that appear in the ($\tilde{\xi}$-approximation) expression for ε, eq. (4·9). They are energy independent and include $ReL_{\pm\frac{1}{2}}$ and $ImL_{-\frac{1}{2}}$, respectively. They are given in Appendix C in terms of $d^{(\sigma)}_{\lambda\lambda'}(\tilde{\xi})$, eqs. (C·46) to (C·52). We can say:

$$P_{T\parallel}=O(1/\tilde{\xi}). \qquad (4\cdot23)$$

This transverse polarization provides a means of detecting the imaginary parts of the coupling constants since the terms involving the real parts goes to zero as zero energy. The $(1/\tilde{\xi})$ corrections to (4·23) are given in Appendix C.

Transverse β polarization perpendicular to the plane of β and γ: This is the distribution which is odd under time reversal (type E of Table 1).

$$P_{T\perp}=\frac{N(+y)-N(-y)}{N(+y)+N(-y)}, \qquad (4\cdot24)$$

where $N(+y)$ and $N(-y)$ are the number of β-rays with their spin parallel and anti-parallel to the y-axis in Fig. 1. The polarization is

$$P_{T\perp}=\frac{E^{(2)}\sin\theta\cos\theta}{C+B^{(2)}(\frac{3}{2}\cos^2\theta-\frac{1}{2})}. \qquad (4\cdot25)$$

The term involving imaginary parts of the coupling constants K_{jk} occurs in $E^{(2)}$ to order $\tilde{\xi}$ with the coefficient $(\alpha Z)^2$. The energy dependence is of the form:

$$P_{T\perp}=\sin\theta\cos\theta\{\tfrac{3}{2}\lambda_9(\alpha Z)^2(I_3/NC)+\tfrac{9}{8}\lambda_8\alpha Z(p/W)(R_3/NC)\}. \tag{4·26}$$

This transverse polarization again strongly resembles the simple directional correlation, $\mathcal{E}$. The R_3 and I_3 are exactly the same quantities that appear in the ($\tilde{\xi}$-approximation) expression for $\mathcal{E}$, eq. (4·9). They are given in Appendix C in terms of $e^{(2)}_{\lambda\lambda'}(\tilde{\xi})$, eqs. (C·55) to (C·61). It is seen that

$$P_{T\perp}=O(\alpha Z/\tilde{\xi}). \tag{4·27}$$

However, for the lighter nuclei, where both these terms become fairly small because of the coefficient αZ, it may be necessary to look for the $(1/\tilde{\xi})$ correction terms, which have no αZ coefficient. This additional contribution to $P_{T\perp}$ among many other different energy terms is

$$\sin\theta\cos\theta\quad(p^2/W)(I_6/NC). \tag{4·28}$$

The order of this term is

$$I_6/NC=O(1/\tilde{\xi}^2). \tag{4·29}$$

Thus, as $\alpha Z\to 0$, $P_{T\perp}$ becomes equal to (4·28) and will be associated only with the imaginary parts of the coupling constants as shown by Curtis and Lewis.[26] Although this transverse polarization is an odd distribution with respect to transformation T, (1·3), it is seen that the possible detection of time reversal violation is made difficult by the relatively large terms involving real parts of coupling constants. This term would have to be separated out by its energy dependence. Indeed it seems that this distribution is less useful for attempts to detect time reversal violation than β-γ correlation or transverse polarization in the plane of the β and γ.

Transverse β polarization perpendicular to the plane of β and γ associated with circular polarization of the γ: This odd distribution with respect to time reversal involves measurement of both β and γ polarizations. It seems almost impossible to do at present. Our conclusion is that even if we would do such an experiment we could not draw any conclusions about the time reversal assumption, because, as will be shown, the contribution from the real part of the coupling constants is much larger than that from the imaginary parts.

We define this quantity analogously to (4·3) and (4·24),

$$P^{P(T)}_{Ty}=\frac{(N(+R)-N(-R))-(N(+L)-N(-L))}{(N(+R)+N(-R))+(N(+L)+N(-L))}, \tag{4·30}$$

where $N(+R)$ et al. are the numbers of β-ray with spin parallel $(+)$ to y-axis and γ-ray with right circular polarizarion (R) and so on. Then from general expression (2·13) and (2·21),

$$P^{P}_{Ty}=-\frac{E^{(1)}+E^{(3)}(\frac{3}{2}\cos^2\theta-\frac{1}{2})}{CN}\sin\theta, \tag{4·31}$$

where $CN=C+B^{(2)}(\frac{3}{2}\cos^2\theta-\frac{1}{2})$. We find in the ξ-approximation

$$P_{Ty}^{\prime}=-\frac{1}{CN}\sin\theta\left[\alpha Z\frac{1}{W}R_7+j_{10}\frac{p}{W}I_7\right], \tag{4·32}$$

where

$$R_7=\sqrt{2}\,(-G_{11}(2)\,y^{(2)}+2G_{01}(2)\,ReVY^*), \tag{4·33}$$

$$I_7=\{-\tfrac{2}{3}G_{01}(2)\,Im[(x_0(-2)-\tfrac{1}{2}u_0(-2))V^*]-\tfrac{1}{3}G_{11}(2)\,Im[(x_0(-2)$$
$$-\tfrac{1}{2}u_0(-2))Y^*]-(\sqrt{15}/6)G_{12}(2)\,Im[z_0(-2)Y^*]\}, \tag{4·34}$$

$$\lambda_{10}=(\sqrt{F_1}/4\sqrt{F_0})\,(j_1+1+j_2)\cos(\theta_2-\theta_1). \tag{4·35}$$

Here F_k and θ_k are defined in Appendix B. Therefore, the order of magnitudes of R- and I_7 which include ReL_{jk} and ImL_{jk}, respectively, are

$$O(R_7/CN)=1,\quad O(I_7/CN)=1\ \xi.$$

We do not present more exact expressions in Appendix C.

For the unique first forbidden transition, there is no transverse polarization except in the last case, i.e. the measurement of polarizations of both β and γ-rays. This expression, of course, includes only absolute values of coupling constants, since no interference between (STP) and (VA) is assumed. We find:

$$P_{Ty}^{\prime}=-\frac{\alpha Z}{CN}\sin\theta\left(\frac{1}{W}\right)\Big[G_{22}(1)\left(\sqrt{\frac{1}{12}}\sqrt{\frac{5}{2}}\right)(q^2|z_0(-1)|^2+p^2\lambda_1|z_0(-2)|^2)$$
$$-\left(\sqrt{\frac{1}{16}}\sqrt{\frac{5}{2}}\right)G_{22}(3)\,p^2\lambda_1(3\cos^2\theta-1)\Big] \tag{4·36}$$

§ 5. Discussion

In the preceding section we have given theoretical results for various observable quantities without specifying the nucleus. We have seen that, roughly speaking, in the non-unique transitions, there are three large observables: the spectrum, longitudinal polarization, and circular polarization; and that there are two types of observables which are relatively small: the β-γ directional correlation and β transverse polarizations. In the unique decay all observables are of order unity $(O(1))$, except no transverse polarization. This situation is reviewed in Table 3, together with information as to which expressions can be found in Appendix C.

We cannot obtain much information about nuclear matrix elements by examining the spectrum or longitudinal polarization in non-unique decays if the ξ-approximation holds exactly within experimental error, see (3·13) and (3·17). In the ξ approximation the circular γ polarization depends on the nuclear matrix elements if $\Delta J=0$, while for $\Delta J=1$ it just depends on the nuclear spins, since $\omega=-\sqrt{2}\,G_{11}(1)$. The smaller effects, the β-γ correlation and β transverse polarizations, do depend directly on the β-decay matrix elements. Through experimental determination of corrections (or lack of them) to the ξ-approximation, however, we may get a great deal more infrormation about nuclear matrix

elements and we learn how accurate the $\tilde{\xi}$-approximation predictions are for other observables. Of course, the original idea for observing corrections to the $\tilde{\xi}$-approximation is in the spectrum shape. Observation of such corrections could be otherwise obtained from (1) deviation of the β-γ correlation coefficient from the energy dependence $\varepsilon \sim W$ for high energy electrons ($W \gg 1$) (see eq. (4·9) etc.); (2) deviation of the longitudinal polarization from $|p, W|$ (see discussion in connection with the RaE type decay in § 3B and § 3C where for high energy electrons the relative deviation is of order $1/W\tilde{\xi}$); (3) deviation from energy independence of the quantity ω in the β-circularly polarized γ correlation (in the $\tilde{\xi}$-approximation ω is constant (see eq. (4·15)).

Table 3

Distribution	Defining Equation	$0(\xi^2)$	$0(\xi)$	$0(1)$
C: Correction factor for the energy spectrum		×	×	×
P_L: Longitudinal polarization	(2–6, 3–1)	×	×	×
ε: β-γ directional correlation	(4–1, 2)		×	×
ω: β- circularly polarized γ	(4–3, 4)	×	×	No
P_L^{γ}: γ- longitudinally polarized β	(4–5)	×	×	No
$P_{T\parallel}$: Transverse polarization in the β-γ plane	(4–21)		×	×
$P_{T\perp}$: Transverse polarization perpendicular to the β-γ plane	(4–24, 25)		×	×
$P^{P}{}_{Ty}$: Transversely polarized β-circularly polarized γ	(4–30, 31)	$\frac{2}{3}4$	No	No

Expressions presented (in Appendix C) are denoted by an ×. The leading term in powers of ξ, and usually the $1/\xi$ corrections to it, are discussed in the text for all these distributions.

In order to test time reversal by observing the β-γ correlation or transverse polarizations it is necessary to measure the energy dependence as discussed in the previous section. In general the contribution of the real parts of coupling constants to these quantities is slightly larger than that of the imaginary parts. Therefore the simplest experiment, the β-γ directional correlation, is preferred. Aside from examination of corrections to the $\tilde{\xi}$-approximation, observation of time reversal violation in the spectrum, longitudinal polarization or circular γ polarization requires absolute determination of the coupling constants. This is not a promising procedure in forbidden decays, except in a special case of strong cancellation like RaE.

As regards the reliability of the theoretical expressions presented in Appendix C, we note that the basic expansion is in terms of the parameter $\tilde{\xi}$. (This expansion is discussed at length in § 3B.) Here $\tilde{\xi}$ denotes either $\alpha Z, 2\rho$ or $\tilde{\xi}'$ (see (3·8)), where ρ is the nuclear radius in electron units ($\approx \frac{1}{2}\alpha A^{1/3}$). The latter quantity is probably more appropriate to describe the $\tilde{\xi}$-expansion in light nuclei. There the expansion may rapidly converge if the energy of the decay is not too large. For heavy nuclei $\alpha Z \cdot 2\rho \approx 1/4\rho$. The terms neglected through $0(1)$ terms is $0(\alpha Z\rho)$, or about 1% for small W_0. The relative error is then at least as small as that when the full expression through $0(1)$ is used. The complexity of the full expression calls for simplification whenever

possible. Thus we have neglected some small corrections (e.g. $(\alpha Z)^2$ terms) to the $0(1)$ terms. (This is discussed in Appendix A. We are careful, however, to present expressions for C and A which are accurate even in case of "*large cancellation*". These terms are needed for RaE). Also, as shown in Table 3, we do not give results for $1/\xi^2$ corrections to the leading term for any β-γ correlation. We feel this is sufficient for the accuracy likely to obtain in these experiments. Furthermore, we propose a series of consistent approximations for $0(\xi)$ or for $0(1)$ terms, in Appendix C, (C·1) to (C·5). These lead to great further simplifications and may be adopted depending on the case and upon the accuracy desired.

Finally, we would like to make a general remark about the first forbidden β-decay. All the large observables (i.e. those of order ξ^2 in Table 3) have counterparts in the allowed transition. The converse is also true. Thus, for every observable that occurs in the allowed transition there is a corresponding expression (of order ξ^2) in the non-unique first forbidden transition, and *vice versa*. The first forbidden expression is obtained (in the ξ-approximation) by using the substitution:

$$\begin{aligned} C_F M_F &\to \gamma_1 V, \\ C_{GT} M_{GT} &\to \gamma_1 Y, \end{aligned} \tag{5·1}$$

where C_F, M_F, C_{GT}, and M_{GT} are the coupling constants and nuclear matrix elements for the Fermi and Gamow-Teller transitions in the allowed transition, respectively. We can for example obtain the result for the β-circularly polarized γ correlation coefficient ω, (4·13), from the allowed transition expression, as shown by Boehm and Wapstra.[7] The preceding statement may be understood from the following similarity between allowed and first forbidden decays. In the allowed transition the essential contribution of the electron wave function is associated with combinations of s-wave radial functions g_{-1} and f_{-1} (where g_κ and f_κ are the usual radial functions associated with small and large components of the spinor, respectively). Thus the transition probabilities involve $L_{00}^{(0)} \sim f_1^2 + g_{-1}^2 \sim 1$ or $L_{00}^{(1)} \sim f_1 g_{-1} \sim p W$ (c.f. Appendix B). In the non-unique first forbidden decay the main contribution involves the same functions for the relativistic nuclear matrix elements, i.e. $\int \alpha$ and $\int \gamma_5$, while it is given by combinations of other radial functions g_{+1} and f_{-1} for the non-relativistic matrix elements; i.e. the quantities $M_{00}^{(0)} \sim (1/r^2)(f_{-1}^2 + g_1^2) \sim (\alpha Z/2\rho)^2$ and $M_{00}^{(1)} \sim (1/r^2) f_{-1} g_1 \sim (\alpha Z/2\rho) p W$ occur instead of the L's above; and so on. The correspondences (5·1) are readily established by direct comparison of these terms.

Allowed transition experiments suggested for testing general properties of the β-decay interaction, can thus be carried through for first forbidden decay. In order to test the time reversal assumption directly, for example, the β-γ angular correlation from oriented nuclei, $(\boldsymbol{J}\cdot\boldsymbol{p}\times\boldsymbol{k})(\boldsymbol{J}\cdot\boldsymbol{k})^{2n+1}$ can be used.[11,12] Concerning to the question which combination of (STP) or (VA) is responsible for the β-decay, the direct way is to observe the nuclear recoil as mentioned in § 1, say eq. (1·1). One of the experiments is to observe the β-neutrino angular correlation, which has the form:

$$N_{e\nu} \sim 1 + \varepsilon_\nu (p/W) \cos\theta_{e\nu}, \tag{5·2}$$

where $\theta_{e\nu}$ is the angle between the β-ray and neutrino directions. The well-known allowed transition expression for ε_ν can be translated into that for the non-unique forbidden transition, in the ξ-approximation, yielding

$$\varepsilon_\nu = \frac{\mp|V|^2 \pm \frac{1}{3}|Y|^2}{|V|^2+|Y|^2} . \tag{5·3}$$

Here the upper and lower signs correspond to the (S, T, P) and (V, A) combinations respectively. (Of course, in the unique forbidden transition, we can not expect such correspondence, because there is no contribution from rank $\lambda=2$ in the allowed transition. For example, the β-neutrino angular correlation is given by

$$\varepsilon_\nu = \pm\frac{p^2+q^2+3pq\cos\theta_{e\nu}}{5\mp pq(p/W)} , \tag{5·4}$$

where the upper and lower signs mean the T and A interactions.[37]) However, these experiments, seemingly equivalent to those in the allowed transition, are probably not very useful to test general properties of the β decay, because the theoretical expressions are not as accurate as in the allowed case, and the matrix element combinations involved are less simple.

Acknowledgement

The authors would like to express their sincere thanks to Professors E. J. Konopinski and M. E. Rose for valuable discussions and to many authors who have sent us their preprints. One of us (T. K.) is deeply indebted to Indiana University for her hospitality.

Appendix A. Contribution of the third forbidden transition

We have presented our discussion above in terms of the usual 6 types of nuclear matrix elements shown in Table 2. We should, however, consider two additional correction terms, especially in case large cancellation takes place as in RaE. One of them arises from more detailed consideration of the nuclear finite size effect. We will discuss it in Appendix B. The other one is conrributions of the so-called third forbidden operators to first forbidden transitions.

In order to see the order of the contribution from the so-called higher order forbidden transition, it is convenient to introduce the following tensor operator, which operates on the nuclear wave functions:

$$T^m_{\lambda Ls}(\boldsymbol{r}\cdot\sigma) = \sum' C(Ls\lambda\,;\,\mu_1\mu_2)\, r^L Y^{\mu_1}_L(\boldsymbol{r})\, Y^{\mu_2}_s(\sigma) , \tag{A·1}$$

where $m=\mu_1+\mu_2$ and C is a Clebsh-Goldan coefficient, $\boldsymbol{r}$ a coordinate vector and $Y^{\mu_2}_s(\sigma)$ stands for the spherical function of the 4×4 Pauli spin operator, σ; for example, $Y^0_0(\sigma)=1/\sqrt{4\pi}$, $Y^1_1(\sigma)=-(\sigma_x+i\sigma_y)\sqrt{3/8\pi}$. The possible values of s are, of course, 0 and 1. In the first forbidden transition, the initial and final nuclei have different parity. Therefore, if we assume parity conservation for the strong interaction, we have two different kinds of operators:

$$(\beta)\, T^{m}_{\lambda L s}(\boldsymbol{r}, \sigma) \qquad \text{for odd L,} \qquad (A\cdot 2)$$

and

$$(-1)^{1-s}(\beta)\,(-i\gamma_5)\, T_{\lambda L s}(\boldsymbol{r}, \sigma) \quad \text{for even L,} \qquad (A\cdot 3)$$

where (β) stands for the β-matrix for the (S, T, P) combination and the unity for the (V, A) combination. The imaginary parameter, i, is introduced to make the nuclear matrix elements real numbers (if the common phase factor is omitted). The usual nuclear operators in "cartesian notation" as listed in Table 2, (1, x, y, and z), correspond to $i\gamma_5 T_{000}$, T_{011}, T_{110} $-i\gamma_5 T_{101}$, T_{111} and T_{211} in that order. For example $4\pi(\beta)\, T_{110}(\boldsymbol{r}, \sigma) = \sqrt{3}\, x = \sqrt{3}\,(\beta)\boldsymbol{r}$ and $4\pi(-i\gamma_5)(\beta)\, T_{101} = \sqrt{3}\, y = \sqrt{3}\, i(\beta)\boldsymbol{\alpha}$. Thus, it is clear that the usual nuclear matrix elements include all terms with $L=0$ and $L=1$. This is how we define the "first forbidden transition". We omitted higher L-values, for example T_{121}, T_{220}, T_{221}, T_{231} and so on. These types are called "third forbidden", i.e. $L=2$, 3 and $\lambda=1$, 2. (We are not interested in larger λ's here). We wish to discuss the contribution from these third forbidden operators.

Table 4

λ	Operator	Nuclear Matrix Element
1	$\phi = -i(\beta)\gamma_5 4\pi T_{121}(\boldsymbol{r}, \sigma)$	$\rho^{\circ}\xi' \eta \phi_0(\kappa) = C_{V(T)} \int \phi R(\kappa)$
2	$\psi = i(\beta)\gamma_5 4\pi T_{220}(\boldsymbol{r}, \sigma)$	$\rho^{\circ}\xi' \eta \psi_0(\kappa) = C_{A(P)} \int \psi R(\kappa)$
2	$\chi = -i(\beta)\gamma_5 4\pi T_{221}(\boldsymbol{r}, \sigma)$	$\rho^{\circ}\xi' \eta \chi_0(\kappa) = C_{T(T)} \int \chi R(\kappa)$

The additional nuclear matrix elements. In the "cartesian notation" which was used in the Konopinski-Uhlenbeck paper[25], there are the following correspondence: $T_{220} \sim R_{ij}$, $T_{221} \sim T_{ij}$, and $T_{221} \sim S_{ijk}$.

Let us consider the general character of $T_{\lambda L s}$ and the corresponding lepton wave functions. It is clear from the conditions, $J_0 - J_1 \le \lambda \le J_0 + J_1$ (3·5), that λ corresponds to the total angular momentum of the lepton field, say $\boldsymbol{\lambda} = \boldsymbol{j} - \boldsymbol{j}_\nu$, and L and S represent the total orbital and spin angulsr momenta of the lepton field, respectively, say $\boldsymbol{L} = \boldsymbol{l}(\kappa) + \boldsymbol{l}(\kappa_\nu)$. In general, the large coulomb correction factor, $\xi = \alpha Z\, 2\rho$, comes from the l-th radial wave of β-particle with $j = l - \frac{1}{2}$, while the l-th wave with $j = l + \frac{1}{2}$ has only minor Coulomb corrections. In other words, the radial wave function with $j = l + \frac{1}{2}$ is similar to the l-th wave for free particles, which is expressed by $j_l(r)$, the spherical Bessel function. For this reason, there is no large coulomb factor, ξ, for the unique forbidden transition operator, T_{211}, and the relativistic nuclear operators, $i\gamma_5$ and $i\alpha$. (In the latter case, the s-wave of β-particle with $j = \frac{1}{2}$ gives the main contribution. In the unique forbidden transition, the p-wave of β-particle with $j = 3/2$ contributes the main term with an s-wave of neutrino with $j_\nu = \frac{1}{2}$, because $\lambda = 2 = \boldsymbol{j} + \boldsymbol{j}_\nu$, or *vice versa*.) The main coefficient for the ordinary nuclear matrix elements with $\sigma \cdot \boldsymbol{r}$ and $\sigma \times \boldsymbol{r}$ is given by the p-wave of β-ray with $j = \frac{1}{2}$. Therefore, the large Coulomb correction factor, ξ, appears in these cases.

Let us consider a typical example with even L, say T_{121} for the (V, A) combination. In the cartesian notation, T_{121} is expressed by

$$4\pi(-i\gamma_5)T_{121}{}^{m}=-(3\sqrt{3}\,i/\sqrt{2})\{(\boldsymbol{\alpha}\cdot\boldsymbol{r})r_m-\tfrac{1}{3}r^2\alpha_m\}. \qquad (A\cdot 4)$$

In analogy with the definition of the nuclear parameters of Table 2, we can define the following ratio

$$\tilde{\xi}'\rho^2\gamma_l\phi_n(\kappa)=C_{V(J')}\int(-i)4\pi\gamma_5({}_l\beta)T_{121}{}^{m}R(\kappa)(r/\rho)^n. \qquad (A\cdot 5)$$

Using this definition, $\phi_n(\kappa)$ is the same order as $\gamma_n(\kappa)$. In defining the ordinary term $\gamma_0(\kappa)$ (which is derived from $(-i\gamma_5)({}_l\beta)T_{101}$), we introduced (Table 2) the coefficient $\tilde{\xi}'$ $(3\cdot 8)$ in order that $\gamma_0(\kappa)$ be of order unity, i.e. $(\int\alpha/\tilde{\xi}'\int\boldsymbol{r})=O(1)$. In the matrix element, $\gamma_0(\kappa)$ appears with, at most, this coefficient, i.e. as $\tilde{\xi}'\gamma_0(\kappa)$. In the case of $\phi_0(\kappa)$ where $L=2$, it appears in the matrix element with, at most, the coefficient $\rho\tilde{\xi}'\alpha Z/2$, because the coefficient arises from the contribution of the p-wave of the β-particle with $J=\frac{1}{2}$ which has the large Coulomb correction factor $\tilde{\xi}$. Strictly speaking, this coefficient can always be written in the form

$$\delta=\rho\tilde{\xi}'\alpha Z/(1+\gamma_1), \qquad (A\cdot 6)$$

for this type of higher order nuclear matrix element with L even. Although we do not have any suitable way of fixing the exact numerical value of $\tilde{\xi}'$ at present, we may say that δ is of order 1/10. (See footnote following $(3\cdot 8)$.) In the transition probability, this type of nuclear matrix element interferes with the ordinary nuclear matrix element which has, at most, a coefficient of order $\tilde{\xi}$. (We are speaking about the non-unique decay). Therefore, the largest contribution to the transition probability associated with the third forbidden transition is of order $\delta\tilde{\xi}$. Then we classify this term as a $\tilde{\xi}$-order term in the $\tilde{\xi}$-expansion in Appendix C.

There are of course smaller contributions resulting from third forbidden-first forbidden interference. Terms in the transition probability of order δ we can classify for simplicity with the unit order terms in the $\tilde{\xi}$ expansion. Actually we keep these small corrections in our expressions in Appendix C only for the case of spectrum and longitudinal polarization which are of interest in the RaE case. The pure third forbidden transition is of order δ^2. This we should neglect, although we include this term sometimes in Appendix C to simplify the final expression. (c.f. $(B\cdot 10)$). (It is of order $\delta^2,\tilde{\xi}^2$ smaller than the leading term in the non-unique first forbidden transition probability.) The discussion for T_{220} and T_{221} is the same as just made for T_{121}.

A discussion similar to that for T_{121} can be made for the nuclear parameter, $\gamma_2(\kappa)$, which, using the notation in Table 2, represents the contribution of $\int\alpha(r/\rho)^2$. Although this term is classed as first forbidden $(L=0)$, the coefficient is given by the p-wave of both β-particle and neutrino. Therefore, the largest coefficient for $\gamma_2(\kappa)$ is of order δ, while that for $\gamma_0(\kappa)$ is of order $\tilde{\xi}'$ (in the matrix element). Therefore, the classification for $\gamma_2(\kappa)$ is the same as that for $\phi_0(\kappa)$. Terms which include $\phi_2(\kappa)$ or $\gamma_4(\kappa)$ and higher values of L have coefficients of order $\rho^2\delta$ or smaller. These terms are neglected.

Next let us consider the contribution from the higher rank nuclear matrix elements, say T_{321}, as an example. In order to get $\lambda=3$ and $L=2$, it is necessary, for example,

to have both p-wave contributions from β-particle with $j=3/2$ and neutrino with $j=3/2$. Thus, the main coefficient for such a high rank nuclear matrix element has no large Coulomb factor ξ. In the matrix element the T_{321} term appears with order of magnitude $(\xi' \rho)\rho$. In the transition probability, the interference of this term with the leading first forbidden term, whose order is ξ, is of order δ (A·6). We may again classify this term with the unit order term in the ξ-expansion. Other higher rank nuclear matrix elements are completely neglected, because their contribution is of order $\alpha Z \cdot \delta$. Thus we have an additional nuclear parameter which is defined as follows:

$$\rho^2 \xi' \eta \zeta_0(\kappa) = C_{V(T)} \int (-i\gamma_5)(\beta) 4\pi T_{321} R(\kappa). \tag{A·7}$$

This term will make contributions like $\rho(\xi' \rho) Re[\zeta_1(\kappa) Y^*]$, to the transition probability, as is clear from the above argument. Such contributions, of order δ, but proportional to Y are neglected in this paper since they become small in the interesting case when Y is small. This neglect will be discussed in connection with other nuclear matrix elements with odd L.

As an example for the operator $T_{\lambda L s}$ with odd L, let us consider T_{231}. We define a quantity analogous to $\phi(\kappa)$, eq. (A·5),

$$\rho^2 \eta \Omega_0(\kappa) = C_{A(T)} \int (\beta) 4\pi T_{231} R(\kappa). \tag{A·8}$$

According to the similar discussion for even L, the contribution from this type of term with odd L is of order $\alpha Z \rho (1+\gamma_1)$, since there is no ξ'-coefficient in the above definition of $\Omega_0(\kappa)$ (c.f. (A·5)). The same considerations apply to the (first forbidden) terms $w_2(\kappa)$, $u_2(\kappa)$ and $x_2(\kappa)$ of Table 2. Thus, in the transition probability, the largest coefficient for such terms is of order $(\alpha Z/2)(\alpha Z(1+\gamma_1))$. This is essentially the same order as δ. Other similar terms are T_{331} and T_{431}. All these terms will occur in combinations, say, like $\alpha Z \rho Re[\Omega_0 Y^*]$ in the transition probability by the same argument as above. Other similar terms occur in the case of usual first forbidden transition probability terms like say, $\alpha Z \rho Re[x_0(\kappa) Y^*]$ and $\rho^2 |Y|^2$. Besides these, terms like $\rho |Y|^2$, which are formally of order $\alpha Z \xi$, should be considered. If there is a large cancellation as in the RaE case where, $|Y|^2 \approx 1$, the contribution from all these terms becomes only of order (1/100) or less. Then they should be neglected, in accord with Assumption (II). However, if there is no such a large cancellation, these correction terms, except $\rho |Y|^2$, are less by a factor $(\alpha Z/2)^2$ or δ than the unity order terms in the ξ-expansion. In Appendix C, these terms are neglected, resulting in an error of order $\delta \cdot \xi \approx 1\%$ in the transition probability relative to a contribution of order ξ. The terms like $\rho |Y|^2$ are classified with the ξ-order terms. These arguments apply to the non-unique decay.

Now let us consider the three new corrections of Table 4 in more detail. In the cases of the energy spectrum and the longitudinal polarization (i.e. no observation of γ-rays) there is no interference term between the different ranks, λ, of nuclear matrix elements. Therefore, in the case of $\lambda=0$, there is no contribution from the third forbidden transition, but we have one extra nuclear parameter $x_2(\kappa)$. In the case of $\lambda=1$, there are

two additional nuclear matrix elements, say $y_2(\kappa)$ and $\phi_0(\kappa)$. In the unique forbidden transition, the correction factor for the energy spectrum can be written as

$$C_2=(1/12)\cdot(q^2|\mathfrak{Z}(-1)|^2+p^2\lambda_1|\mathfrak{Z}(-2)|^2), \tag{A·9}$$

where

$$\mathfrak{Z}(-k)=z_0(-k)+(2\delta(1+\gamma_1)/\sqrt{15}(k+\gamma_1))(x_0(k)-\sqrt{2}\,\psi_0(k)). \tag{A·10}$$

In the preceding sections, we did not show this complicated formula, because it leads to no essential difference. In all results for the unique forbidden transition, $|\mathfrak{Z}|^2$ should be used instead of $|z_0|^2$. It is worthwhile to note that all the other correction terms to the unique forbidden transition make a contribution of order $\alpha Z\rho$ (say, terms like $\alpha Z\rho|z_0(-1)|^2$).

If the β-γ correlation is measured, the situation is a little more complicated, because there are many interference terms between different ranks of the nuclear matrix elements. In Appendix C, all terms with coefficient δ or δ^2 are neglected as mentioned above except the terms whose coefficient is of order of $\xi\delta$. It should be noted that the $\cos^4\theta$ term in the β-γ correlation is a minor correction term.

In general, the β-γ directional correlation function can be expressed by

$$N=C+\sum_{n\geq 1}B^{(n)}P_n(\boldsymbol{p}\cdot\boldsymbol{k}), \tag{A·11}$$

where $B^{(n)}$ is defined by (2·15) and $P_n(\boldsymbol{p}\cdot\boldsymbol{k})$ is a Legendre function. The possible value of n is restricted by

$$0\leq n\leq 2J_1, \tag{A·12}$$

$$|\lambda-\lambda'|\leq n\leq\lambda+\lambda', \tag{A·13}$$

and in addition

$$|j-j'|\leq n\leq j+j', \tag{A·14}$$

as shown in (2·11). Here j is the total angular momentum of β-particle. As long as we consider only the usual 6 nuclear matrix elements of Table 2 we need not consider the contribution from P_n with $n\geq 4$, since $j\leq\frac{3}{2}$ and $\lambda\leq 2$. Although the new nuclear matrix elements discussed here have the maximum value of $\lambda=3$, the main contribution is still given by the β-ray wave function with $j\leq\frac{3}{2}$. Thus, we may say that the ratio $B^{(4)}/B^{(2)}$ is of order (ρ/ξ).

Appendix B. The finite nuclear size effect

In this pppendix, the basic form of the particle parameter is demonstrated for a few examples, A_1, C_1, and $b_{02}{}^{(2)}$, consistent with assumptions discussed in this paper. In this form the actual β-ray radial functions are included explicity. The more convenient form given in Appendix C is obtained after expansion of the electron radial functions. The expressions shown here illustrate the modifications of previous expressions, such as those of Konopinski and Uhlenbeck,[25] which are necessary to take into account finite nuclear size effects. A_1 and C_1 are the parameter which give the longitudinal polarization and

the shape correction factor in the $0 \rightleftarrows 1$ transition (with parity change), respectively. $b_{02}^{(2)}$ is one of the particle parameters which is necessary to get the β-γ directional correlation in the $\Delta J=0$ transition, where $J_0+J_2 \geq 2$.

In order to define these parameters conveniently, we will use the normalization factor,

$$\gamma_1=2\binom{1+\gamma_1}{2}|\gamma|^2. \tag{B·1}$$

The upper and lower signs of the symbol $(\pm)$ in the following expressions refer to (S, T, P) and (V, A) combinations, respectively. The definitions of nuclear operators, π, ν and others are defined in Tables 2 and 4.

$$\begin{aligned}
\gamma_1 A_1=\{&L_{VV(SS)}[\pm M_{00}^{\prime(+)}(x, x)+(2/3)qN_{00}^{\prime(+)}(x, x)\pm(1/3)q^2L_{00}^{\prime(+)}(x, x)\\
&\pm 2L_{11}^{\prime(+)}(x, x)]\\
&+L_{AA(TT)}[\pm M_{00}^{\prime(+)}(u, u)-(2/3)qN_{00}^{\prime(+)}(u, u)\pm(1/6)q^2L_{00}^{\prime(+)}(u, u)\\
&\pm(1/2)L_{11}^{\prime(+)}(u, u)]\\
&+L_{VV(TT)}[\pm L_{00}^{\prime(+)}(y, y)+(2/9)qN_{00}^{\prime(+)}(yr^2, y)\\
&+(4/9)(2/3)^{1/2}qN_{00}^{\prime(-)}(\phi, y)]\\
&+2Re[L_{AV(TT)}][-(\pm)N_{00}^{\prime(+)}(u, y)+(1/3)qL_{00}^{\prime(+)}(u, y)\\
&-(1/9)qM_{00}^{\prime(+)}(u, yr^2)\pm(1/9)q^2N_{00}^{\prime(+)}(yr^2, u)\\
&-(2/9)(2/3)^{1/2}qM_{00}^{\prime(+)}(u, \phi)\pm(1/18)(2/3)^{1/2}q^2N_{00}^{\prime(+)}(\phi, u)\\
&\pm(1/6)^{1/2}N_{11}^{\prime(+)}(\phi, u)]\\
&+2Re[L_{VV(ST)}][\pm N_{00}^{\prime(+)}(x, y)+(1/3)qL_{00}^{\prime(+)}(u, y)+(1/9)qM_{00}^{\prime(+)}(x, yr^2)\\
&-(\pm 1/9)q^2N_{00}^{\prime(+)}(yr^2, x)+(2/9)(2/3)^{1/2}qM_{00}^{\prime(+)}(x, \phi)\\
&\pm(1/9)(2/3)^{1/2}q^2N_{00}^{\prime(+)}(\phi, x)\pm(2/3)^{1/2}N_{11}^{\prime(+)}(\phi, x)]\\
&+2Re[L_{VA(ST)}][-(\pm)M_{00}^{\prime(+)}(x, u)+(1/3)qN_{00}^{\prime(+)}(x, u)\\
&-(1/3)qN_{00}^{\prime(+)}(u, x)\pm L_{11}^{\prime(-)}(x, u)]\\
&-2Im[L_{AV(TT)}][-(\pm)J_{00}^{\prime(+)}(u, y)-(\pm 1/9)q^2J_{00}^{\prime(+)}(yr^2, u)\\
&-(\pm 1/18)(2/3)^{1/2}q^2J_{00}^{\prime(+)}(\phi, u)\\
&-(\pm 1/2)(2/3)^{1/2}J_{11}^{\prime(+)}(\phi, u)]\\
&-2Im[L_{VV(ST)}][\pm J_{00}^{\prime(+)}(x, y)\pm(1/9)q^2J_{00}^{\prime(+)}(yr^2, x)\\
&-(\pm 1/9)(2/3)^{1/2}q^2J_{00}^{\prime(+)}(\phi, x)\\
&-(\pm)(2/3)^{1/2}J_{11}^{\prime(+)}(\phi, x)]\\
&-2Im[L_{VA(ST)}](1/3)q[J_{00}^{\prime(+)}(x, u)+J_{00}^{\prime(+)}(u, x)]\}
\end{aligned} \tag{B·2}$$

$\gamma_1 C_1$ is found easily from $\gamma_1 A_1$ by the following substitutions:

$$L_{ij} \to \pm K_{ij}, \quad (L', M', N', J') \to (L, M, N, J)$$

It is clear from the definitions H, J and K that $H^{(1)}_{kk} = J_{kk} = K^{(1)}_{kk} = 0$. Therefore, all $Im[K_{ij}]$ automatically disappear in $\eta_1 C_1$.

$$\begin{aligned}
\eta_1 b^{(2)}_{02} = (3/2)^{1/2} \{ & 2Re[K_{AA(PT)}][L'^{(+)}_{01}(v, z) - (\pm 1/3) q N'^{(+)}_{01}(vr^2, z) \\
& \pm (1/15) q N'^{(+)}_{10}(zr^2, v)] \\
& - 2K_{AA(TT)}[N'^{(+)}_{01}(w, z) \pm (1/3) q L'^{(+)}_{01}(w, z) \pm (1/15) q M'^{(+)}_{01}(w, zr^2)] \\
& + 2Im[K_{AA(PT)}][-H'^{(+)}_{01}(v, z) \pm (1/3) q J'^{(+)}_{01}(vr^2, z) \pm (1/15) q J'^{(+)}_{10}(zr^2, v)]\} \\
+ (2/5)^{1/2} \{ & -2^{3/2} Re[K_{AA(TP)}] M'^{(+)}_{01}(w, \psi') + 2^{3/2} K_{AA(PT)} N'^{(+)}_{10}(\psi', v) \\
& + 2Re[K_{AV(TT)}] M'^{(+)}_{01}(w, \chi) - 2Re[K_{AV(PT)}] N'^{(+)}_{10}(\chi, v) \\
& + 2^{3/2} Im[K_{AA(TP)}] K'^{(+)}_{01}(w, \psi) \\
& - 2Im[K_{AV(TT)}] K'^{(+)}_{01}(w, \chi) - 2Im[K_{AV(PT)}] J'^{(+)}_{10}(\chi, v)\}. \qquad (B\cdot 3)
\end{aligned}$$

The definitions of L, M, N, H, J and K are given below as certain combinations of β-ray radial wave functions, $f_\varkappa$ and $g_\varkappa$. As we assume that the strong and electromagnetic interactions are invariant under time reversal, (Assumption III), all these L, M, N, H, J, and K are real.

$$L^{(\prime)(\pm)}_{k-1,m-1}(O_i, O_j) = \varepsilon[L^{(\prime)g}_{k-1,m-1} \pm L^{(\prime)f}_{k-1,m-1}], \qquad (B\cdot 4)$$

$$H^{(\prime)(\pm)}_{k-1,m-1}(O_i, O_j) = \varepsilon[H^{(\prime)g}_{k-1,m-1} \pm H^{(\prime)f}_{k-1,m-1}], \qquad (B\cdot 5)$$

where $\varepsilon = +1$ for the unprimed cases and $\varepsilon = -1$ for the primed cases.

$$\begin{bmatrix} L^g \\ H^g \end{bmatrix} = \begin{bmatrix} \cos(\Delta_{-k} - \Delta_{-m}) \\ \sin(\Delta_{-k} - \Delta_{-m}) \end{bmatrix} \left(\int O_i g_{-k} r^{-k+1}\right) \left(\int O_j g_{-m} r^{-m+1}\right) * (2p^2 F_0)^{-1}, \qquad (B\cdot 6)$$

$$\begin{bmatrix} L_f \\ H^f \end{bmatrix} = \begin{bmatrix} \cos(\Delta_k - \Delta_m) \\ \sin(\Delta_k - \Delta_m) \end{bmatrix} \left(\int O_i f_k r^{-k+1}\right) \left(\int O_j f_m r^{-m+1}\right) * (2p^2 F_0)^{-1}, \qquad (B\cdot 7)$$

$$\begin{bmatrix} L'^g \\ H'^g \end{bmatrix} = \begin{bmatrix} \cos(\Delta_{-k} - \Delta_{+m}) \\ \sin(\Delta_{-k} - \Delta_{+m}) \end{bmatrix} \left(\int O_i g_{-k} r^{-k+1}\right) \left(\int O_j f_m r^{-m+1}\right) * (2p^2 F_0)^{-1}, \qquad (B\cdot 8)$$

$$\begin{bmatrix} L'^f \\ H'^f \end{bmatrix} = \begin{bmatrix} \cos(\Delta_k - \Delta_{-m}) \\ \sin(\Delta_k - \Delta_{-m}) \end{bmatrix} \left(\int O_i f_k r^{-k+1}\right) \left(\int O_j g_{-k} r^{-m+1}\right) * (2p^2 F_0)^{-1}, \qquad (B\cdot 9)$$

$$N^{(\prime)(\pm)}_{k-1,m-1}(O_i, O_j) = [N^{(\prime)g} \mp N^{(\prime)f}], \qquad (B\cdot 10)$$

$$J^{(\prime)(\pm)}_{k-1,m-1}(O_i, O_j) = [J^{(\prime)g} \pm J^{(\prime)f}], \qquad (B\cdot 11)$$

$$\begin{bmatrix} N^g \\ J^g \end{bmatrix} = \begin{bmatrix} \cos(\Delta_{-k} - \Delta_{-m}) \\ \sin(\Delta_{-k} - \Delta_{-m}) \end{bmatrix} \left(\int O_i f_{-k} r^{-k}\right) \left(\int O_j g_{-m} r^{-m+1}\right) * (2p^2 F_0)^{-1}, \qquad (B\cdot 12)$$

$$\begin{bmatrix} N^f \\ J^f \end{bmatrix} = \begin{bmatrix} \cos(\Delta_k - \Delta_m) \\ \sin(\Delta_k - \Delta_m) \end{bmatrix} \left(\int O_i g_k r^{-k}\right) \left(\int O_j f_m r^{-m+1}\right) * (2p^2 F_0)^{-1}, \qquad (B\cdot 13)$$

$$\begin{bmatrix} N'^g \\ J'^g \end{bmatrix} = \begin{bmatrix} \cos(\Delta_k - \Delta_{-m}) \\ \sin(\Delta_k - \Delta_{-m}) \end{bmatrix} \left(\int O_i g_k r^{-k}\right) \left(\int O_j g_{-m} r^{-m-1}\right)^* (2p^2 F_0)^{-1}, \tag{B·14}$$

$$\begin{bmatrix} N'^f \\ J'^f \end{bmatrix} = \begin{bmatrix} \cos(\Delta_{-k} - \Delta_m) \\ \sin(\Delta_{-k} - \Delta_m) \end{bmatrix} \left(\int O_i f_{-k} r^{-k}\right) \left(\int O_j f_m r^{-m-1}\right)^* (2p^2 F_0)^{-1}, \tag{B·15}$$

$$M^{(\prime)(\pm)}_{k-1,m-1}(O_i, O_j) = [M^{(\prime)g} \pm M^{(\prime)f}], \tag{B·16}$$

$$K^{(\prime)(\pm)}_{k-1,m-1}(O_i, O_j) = [K^{(\prime)g} \pm K^{(\prime)f}], \tag{B·17}$$

$$\begin{bmatrix} M^g \\ K^g \end{bmatrix} = \begin{bmatrix} \cos(\Delta_k - \Delta_m) \\ \sin(\Delta_k - \Delta_m) \end{bmatrix} \left(\int O_i g_k r^{-k}\right) \left(\int O_j g_m r^{-m}\right)^* (2p^2 F_0)^{-1}, \tag{B·18}$$

$$\begin{bmatrix} M^f \\ K^f \end{bmatrix} = \begin{bmatrix} \cos(\Delta_{-k} - \Delta_{-m}) \\ \sin(\Delta_{-k} - \Delta_{-m}) \end{bmatrix} \left(\int O_i f_{-k} r^{-k}\right) \left(\int O_j f_{-m} r^{-m}\right)^* (2p^2 F_0)^{-1}, \tag{B·19}$$

$$\begin{bmatrix} M'^g \\ K'^g \end{bmatrix} = \begin{bmatrix} \cos(\Delta_k - \Delta_{-m}) \\ \sin(\Delta_k - \Delta_{-m}) \end{bmatrix} \left(\int O_i g_k r^{-k}\right) \left(\int O_j f_{-m} r^{-m}\right) ({}^*2p^2 F_0)^{-1}, \tag{B·20}$$

$$\begin{bmatrix} M'^f \\ K'^f \end{bmatrix} = \begin{bmatrix} \cos(\Delta_{-k} - \Delta_m) \\ \sin(\Delta_{-k} - \Delta_m) \end{bmatrix} \left(\int O_i f_{-k} r^{-k}\right) \left(\int O_j g_{-m} r^{-m}\right) ({}^*2p^2 F_0)^{-1}, \tag{B·21}$$

The Coulomb phase shift, Δ_κ, is defined by

$$\Delta_\kappa = \tfrac{1}{2}\arg.\left(\frac{-\kappa + i\cdot(\alpha Z/p)}{\gamma_k + i(\alpha ZW/p)}\right) - \theta_k - \tfrac{1}{2}\pi(1 + (\kappa/k)), \tag{B·22}$$

where

$$|\kappa| = k \quad \text{or} \quad m, \tag{B·23}$$

$$\theta_k = \arg.\Gamma(\gamma_k + i\alpha ZW/p) + \tfrac{1}{2}\pi(\gamma_k - k). \tag{B·24}$$

The $\arg\Gamma$ has been tabulated.[38)]
The definition of Fermi function is

$$F_{k-1}(Z, W) = \left(\frac{(2k)!}{(k-1)!}\right)^2 (2p\rho)^{2(\gamma_k - k)} e^{\pi\alpha ZW/p} \frac{|\Gamma(\gamma_k + i\alpha ZW/p)|^2}{(\Gamma(1 + 2\gamma_k))^2}. \tag{B·25}$$

O_i's stand for the nuclear operators multiplied by r^n. For example, $O_{r(\alpha)} = \alpha r^n = (\beta) \mathbf{r}\, r^n$.

These more complicated definitions of L, M and so on are introduced to take into account the finite nuclear size effect. In order to show the energy dependence explicity, two ratios, $R_g(\kappa)$ and $R_f(\kappa)$ are introduced in (3·2a) of § 3A, which are ratios of the large (g_κ) and small (f_κ) components of the actual β-ray wave function to the corresponding components of the point charge wave function at the nuclear surface.

If we assume that

$$R_g(\kappa) = R_f(\kappa) = 1, \tag{3·3a}$$

we get the usual definition of Konopinski and Uhlenbeck and others.[25,11,26] For example,

$$L^{(+)}_{k-1,k-1}(O_i, O_j) \Rightarrow \left(\int O_i\right)\left(\int O_j\right)^* L_{k-1}, \tag{B·27}$$

$$L^{(-)}_{k-1,k-1}(O_i, O_j) \Rightarrow \left(\int O_i\right)\left(\int O_j\right)^* L^-_{k-1} \quad (\text{or } P_{k-1}). \tag{B·28}$$

But this approximation includes in some cases perhaps 20% errors, as shown in (3·3b).

In order to avoid these errors, we defined the nuclear parameters $u(\kappa)$, $v(\kappa), \cdots, z(\kappa)$ by using (only) the leading term, $R(\kappa)$, of these R_g and R_f. (c.f. (3·4) and Table 2.) But as mentioned in § 3A, the neglected terms $({}_\rho R^{(1)}_g(\kappa)$ and ${}_\rho R^{(1)}_f(\kappa))$ of $R_g(\kappa)$ and $R_f(\kappa)$ are of order ρ. Therefore, in the ξ expansion, $R^{(1)}(k)$ and $R^{(1)}(-k)$ yield corrections which should be classified in the $1/\xi$ correction term to the leading term in the ξ expansion. In order to take into account these corrections, it is convenient to introduce two ratios:

$$(1/\alpha Z)\, S(k) = R^{(1)}_g(k) - R^{(1)}_f(k), \tag{B·29a}$$

$$\alpha Z S(-k) = R^{(1)}_g(-k) - R^{(1)}_f(-k), \tag{B·29b}$$

$$(W/\alpha Z)\, T(k) = R^{(1)}_g(k) + R^{(1)}_f(k), \tag{B·30a}$$

$$(\alpha Z W)\, T(-k) = R^{(1)}_g(-k) + R^{(1)}_f(-k), \tag{B·30b}$$

$S(\kappa)$ and $T(\kappa)$ are energy independent to $(\alpha Z)^2$ order, because the original Dirac equation is the same for g_κ and $f_{-\kappa}$, except for the sign of mass term. In the case of $\kappa = k$, for the uniform charge distribution,

$$S(k) \simeq T(k) \approx \frac{4k}{(2k+1)^2}\left[1 - k + \frac{k(2k+1)}{(2k+3)}\left(\frac{r}{\rho}\right)^2\right] + O(\alpha Z^2) \tag{B·31}$$

neglecting $(\alpha Z)^2$-terms.*

Thus, two new nuclear parameters for every nuclear operator O_i should be defined in a similar way to that in Table 2. For example,

$$\eta\, u_0{}^S(\kappa) = C_{A(T)} \int u S(\kappa), \tag{B·32}$$

$$\eta\, u_0{}^T(\kappa) = C_{A(T)} \int u T(\kappa). \tag{B·33}$$

The order of magnitude of $u_0{}^S(k)$ may be understood by using (B·31) and the following rough approximation (for the uniform charge distribution):

$$\frac{u_0{}^S(+1)}{u_0(-1)} \approx \frac{u_0{}^T(+1)}{u_0(-1)} \approx \left[\frac{S(+1)}{R(-1)}\right]_{r=\rho} \approx \frac{4}{15}. \tag{B·34}$$

A similar numerical ratio is obtained for the case of $u_0{}^S(-k)$ and $u_0{}^T(-k)$.

---

* In the case of the surface charge distribution,

$$S(k) \simeq T(k) \sim 4k(1+k)(2k+1)^{-2}.$$

Appendix C. Particle parameters

In this Appendix, all particle parameters are given. The theoretical predictions for the transition probabilities are obtained on substitution in (2·15) and thence into (2·13) or (2·19). For convenience the particle parameters are expressed in form of the ξ-expansion (§ 3B).

Now let us consider an example for the case of the energy spectrum with $3^- \to 2^-$ decay. Since $J_1+J_0=5$, we should consider both $\lambda=1$ and 2. In the ξ-expansion, the leading terms $C_1(\xi^2)$, (C·14), and $C_2(1)$, (C·21), for $\lambda=1$ and 2, are of order ξ^2 and 1, respectively. Therefore the spectrum shape correction factor C is expressed by $C_1(\xi^2)$, which is energy independent. In order to consider the deviation from this allowed shape, the next correction term to the shape correction factor, is given by $C_1(\xi)$, (C·15). In this correction, we find so many unknown nuclear parameters that the expression may be of little value.

There is, however, a rough approximation by which we can simplify the correction $C_1(\xi)$, i.e. to neglect perhaps 10% to 20% of $C_1(\xi)$. (This corresponds to the so-called $(\alpha Z)^2$-approximation). We can neglect $\delta \simeq \frac{1}{2}\alpha Z\xi' \rho$ (A·6). Thus we set

$$\delta=0. \tag{C·1}$$

A part of this neglect corresponds to omission of the contributions from the third forbidden transition. Simultaneously we should simplify the finite size function $R(\kappa)$ involved in the nuclear parameters. For the uniform charge distribution, evaluating R, T, S at $r=\rho$ (see (3·36) and (B·31)), we can set;

$$\begin{aligned} &R(\kappa)=R(-1), \quad \text{for all negative } \kappa \\ &R(+1)\approx 0.8R(-1), \\ &R(+2)\approx 0.9R(-1), \\ &S(+1)\approx T(+1)\approx 0.3R(-1). \end{aligned} \tag{C·2}$$

We may write approximately, for $u(\kappa)$, for example,

$$\frac{u(\kappa)}{u(-1)} \simeq \frac{R(\kappa)}{R(-1)}\bigg|_{\rho=T} \approx \begin{cases} 1 & \text{for } |\kappa| \neq 1, \\ 0.8 & \text{for } \kappa=1, \end{cases} \tag{C·3}$$

$$u^T(\kappa)=u^S(\kappa)\approx 0. \tag{C·4}$$

Thereby the independence of the $u(\kappa)$ can be eliminated. This approximation is probably fairly good. The approximation depends both on the nuclear charge distribution and strongly on the nuclear radial functions. To be consistent, we should also neglect $(\alpha Z)^2$ terms so that

$$\Gamma_k=\lambda_k=\gamma_k=1. \tag{C·5}$$

Since these approximations were applied to the second term $C_1(\xi)$ in the ξ-expansion, for our example, they are not so unreasonable. The relations (C·1) to (C·5) may be used in this way in many cases. For some other cases, however, the situation is not

quite as good. For the β-γ correlation, for example, the *leading* term, $b^{(2)}_{\lambda\lambda'}(\tilde{\xi})$ is very complicated. It is again advantageous to use (C·1) to (C·5) but the approximation is of course much poorer. Its use obviously depends on the accuracy of the data considered. Sometimes, unfortunately, this simplification clearly cannot be used at all, as for example, the spectrum of the RaE decay.

All results in this Appendix are shown for the β^--decay. The corresponding expressions for the β^+-decay are obtained in the notation used here by making transformations: $Z\to -Z$, $A_\lambda\to -A_\lambda$, $b^{(n)}_{\lambda\lambda'}\to(-1)^n b^{(n)}_{\lambda\lambda'}$, $a^{(2)}_{\lambda\lambda'}\to -a^{(2)}_{\lambda\lambda'}$, $d^{(2)}_{\lambda\lambda'}\to -d^{(2)}_{\lambda\lambda'}$, $e^{(2)}_{\lambda\lambda'}\to +e^{(2)}_{\lambda\lambda'}$, $w_n(\kappa)\to \mp w_n^*(\kappa)$, $v_n(\kappa)\to v_n^*(\kappa)$, $u_n(\kappa)\to \mp u_n^*(\kappa)$, $x_n\to \pm x_n^*(\kappa)$, $y_n(\kappa)\to -y_n^*(\kappa)$, $z_n(\lambda)\to \mp z_n^*(\kappa)$, $\phi_0(\kappa)\to -\phi_0^*(\kappa)$, $\psi_0^*(\kappa)\to\psi(\kappa)$, $\chi_0(\kappa)\to -\chi_0^*(\kappa)$. All upper and lower signs in this Appendix, including the symbol $(\pm q)$, refer to the (S, T, P) and (V, A) combinations, respectively. $(q=W_0-W.)$

(1) Spectrum Shape Correction Factor (C_λ):

(2) Longitudinal Polarization (A_λ):

$$\begin{bmatrix} C_0=C_0(\tilde{\xi}^2)+C_0(\tilde{\xi})+C_0(1), \\ A_0=(-p/W)\{A_0(\tilde{\xi}^2)+A_0(\tilde{\xi})+A_0(1)\}, \end{bmatrix} \tag{3·9}$$

$$\begin{bmatrix} C_0(\tilde{\xi}^2) \\ A_0(\tilde{\xi}^2) \end{bmatrix} = |V|^2, \tag{C·6}$$

$$\begin{bmatrix} C_0(\tilde{\xi}) \\ A_0(\tilde{\xi}) \end{bmatrix} = \{(2W/3)\,Re[(\mathfrak{W}_{+T}-\delta(5\Gamma_1/\tilde{\xi}')V)V^*]-(\pm 2q/3)\,Re[\mathfrak{W}_-V^*] + \begin{bmatrix} -(2\gamma_1/3W)\,Re[(\mathfrak{W}_{+S}+\delta(1+\gamma_1)\Gamma_0(2\tilde{\xi}')^{-1}V)V^*] \\ +\alpha Z(2/3p)\,Im[\mathfrak{W}_{+S}V^*] \end{bmatrix}\}, \tag{C·7}$$

$$\begin{bmatrix} C_0(1) \\ A_0(1) \end{bmatrix} = (1/9)\{W^2|\mathfrak{W}_{+T}|^2+q^2|\mathfrak{W}_-|^2-2(\pm qW)\,Re[\mathfrak{W}_{+T}\mathfrak{W}_-^*] - \begin{bmatrix} \gamma_1^2|\mathfrak{W}_{+T}|^2-(9/4)|w^V-\gamma_1 w^T|^2 \\ |\mathfrak{W}_{+S}|^2 \end{bmatrix} + \begin{bmatrix} 2\gamma_1(\pm q/W)\,Re[\mathfrak{W}_{+S}\mathfrak{W}_-^*] \\ \alpha Z(9W/p)\,I_2 \end{bmatrix}\}, \tag{C·8}$$

Definitions:

$$V=\xi' v_0(-1)+(2\tilde{\xi}/1+\gamma_1)w_0(+1), \tag{3·10}$$

$$2\mathfrak{W}_{+T}=2\Gamma_0 w_0(+1)+3w_0^T, \tag{C·9a}$$

$$2\mathfrak{W}_{+S}=2\gamma_1\Gamma_0 w(+1)+3w_0^S, \tag{C·9b}$$

$$\mathfrak{W}_-=w_0(-1)-\delta v_2(+1), \tag{C·10}$$

$$w^T=\{w_0^T(+1)+\delta(1+\gamma_1)^2 v_0^T(-1)\}/(1+\gamma_1), \tag{C·11a}$$

$$w^S=\{w_0^S(+1)-\delta(1+\gamma_1)^2 v_0^S(-1)\}/(1+\gamma_1), \tag{C·11b}$$

$$I_2=\gamma_1\Gamma_0(6\hat{\xi})^{-1}Im[(w_0^T(+1)-\delta(1+\gamma_1)^2v_0^T(-1))V^*]$$
$$-2\delta(1+\gamma_1)^{-1}\hat{\xi}^{-1}Im[w_0(+1)V^*]+\Gamma_0(\gamma_1-2\alpha^2Z^2)(1+\gamma_1)(6\hat{\xi})^{-1}Im[w^SV^*]$$
$$+\tfrac{1}{2}\delta Im[w_0^T(+1)v_0^{S*}(-1)+w_0^S(+1)v_0^{T*}(-1)], \quad \text{(C·11c)}$$

$$\delta=\alpha Z\delta\hat{\xi}'/(1+\gamma_1), \quad \text{(A·10)}$$

$$\Gamma_0=6/((1+\gamma_1)(2\gamma_1+1)), \quad \gamma_k=\sqrt{k^2-(\alpha Z)^2}, \quad \text{(C·12)}$$

$$\Gamma_1=3(2\gamma_1+3)/5(2\gamma_1+1). \quad \text{(C·13)}$$

$$\begin{bmatrix} C_1=C_1(\hat{\xi}^2)+C_1(\hat{\xi})+C_1(1), \\ A_1=(-p/W)\{A_1(\hat{\xi}^2)+A_1(\hat{\xi})+A_1(1)\}, \end{bmatrix} \quad \text{(C·14)}$$

$$\begin{bmatrix} C_1(\hat{\xi}^2) \\ A_1(\hat{\xi}^2) \end{bmatrix}=|Y|^2,$$

$$\begin{bmatrix} C_1(\hat{\xi}) \\ A_1(\hat{\xi}) \end{bmatrix}=\Big\{-(2W/3)Re[(\mathfrak{U}_{+T}+\delta(5\Gamma_1/\hat{\xi}')Y)Y^*]$$
$$-(\pm 2q/3)Re[\mathfrak{U}_-Y^*]$$
$$+\begin{bmatrix} (2\gamma_1/3W)Re[(\mathfrak{U}_{+S}-\delta((1+\gamma_1)\Gamma_0/2\hat{\xi}')Y)Y^*] \\ -\alpha Z(2/3p)Im[\mathfrak{U}_{+S}Y^*] \end{bmatrix}$$
$$+\begin{bmatrix} 0 \\ -\alpha Z(2/p)(1+\gamma_1)^{-2}\hat{\xi}Im[u_0^S(1)x_0^*(1)+x_0^S(1)u_0^*(1)] \end{bmatrix}\Big\}, \quad \text{(C·15)}$$

$$\begin{bmatrix} C_1(1) \\ A_1(1) \end{bmatrix}=\left(\frac{1}{9}\right)\Big\{W^2|\mathfrak{U}_{-T}|^2+2p^2\lambda_1|\mathfrak{X}(-2)|^2$$
$$+q^2(|\mathfrak{U}(-1)|^2+2|\mathfrak{X}(-1)|^2)$$
$$+2(\pm qW)Re[\mathfrak{U}_-\mathfrak{U}_{+T}]$$
$$-\begin{bmatrix} \gamma_1^2|\mathfrak{U}_{+T}|^2-(9/4)|u^S-\gamma_1u^T|^2 \\ |\mathfrak{U}_{+S}|^2 \end{bmatrix}$$
$$-\begin{bmatrix} 2\gamma_1(\pm q/W)Re[\mathfrak{U}_-\mathfrak{U}_{+S}^*] \\ -\alpha Z(9W/p)\Gamma_0'+\alpha Z(q/p)\Gamma_1' \end{bmatrix}\Big\}, \quad \text{(C·16)}$$

Definitions :

$$Y=\hat{\xi}'y_0(-1)-2\hat{\xi}(1+\gamma_1)^{-1}(u_0(+1)+x_0(+1)), \quad \text{(3·11)}$$

$$2\mathfrak{U}_{+T}=2\Gamma_0(u_0(+1)+x_0(+1))+3u_0^T, \quad \text{(C·17a)}$$

$$2\mathfrak{U}_{+S}=2\gamma_1\Gamma_0(u_0(+1)+x_0(+1))+3u_0^S, \quad \text{(C·17b)}$$

$$\mathfrak{U}_-=u_0(-1)-x_0(-1)-\delta\phi, \quad \text{(C·17c)}$$

$$u_0^T=\{u_0^T(+1)+x_0^T(+1)-\delta(1+\gamma_1)^2y_0^T(-1)\}/(1+\gamma_1), \quad \text{(C·18a)}$$

$$u^S=\{u_0{}^S(+1)+x_0{}^S(+1)+\delta(1+\gamma_1)^2 y_0{}^S(-1)\}/(1+\gamma_1), \tag{C·18b}$$

$$\phi=(2/3)^{3/2}\phi_0(+1)-(1/3)y_2(+1), \tag{C·19}$$

$$\mathfrak{U}(-1)=u_0(-1)+x_0(-1)+\delta y_2(+1), \tag{C·20a}$$

$$\mathfrak{X}(-k)=x_0(-k)-\tfrac{1}{2}u_0(-k)+\delta(1+\gamma_1)\phi_0(k)/\sqrt{6}\,(k+\gamma_k), \tag{C·20b}$$

$$\begin{aligned} I_2'=&\{(1+\gamma_1)\Gamma_0\tilde{\xi}'(6\tilde{\xi})^{-1}Im[(\gamma_1 u^T+(\gamma_1-2\alpha^2Z^2)u^S)y_0{}^*(-1)] \\ &+\delta(1+\gamma_1)^2\Gamma_0(6\tilde{\xi})^{-1}Im[(\gamma_1 y_0{}^T(-1)+(\gamma_1-2\alpha^2Z^2)y_0{}^S(-1))Y^*] \\ &-4\delta(1+\gamma_1)^{-1}\tilde{\xi}^{-1}Im[(u_0(+1)+x_0(+1))Y^*] \\ &+\tfrac{1}{2}\delta(1+\gamma_1)Im[u^S y_0{}^{T*}(-1)+y_0{}^S(-1)u^{T*}]\}, \end{aligned} \tag{C·20c}$$

$$\begin{aligned} I_2''=&\{2\gamma_1\Gamma_0 Im[x_0(1)u_0{}^*(-1)+u_0(1)x_0{}^*(-1)] \\ &+3(1+\gamma_1)^{-1}Im[x_0{}^S(1)v_0{}^*(-1)+u_0{}^S(1)x_0{}^*(-1)] \\ &+3\alpha^2Z^2(1+\gamma_1)^{-1}Im[x_0(1)v_0{}^{S*}(-1)+u_0(1)x_0{}^{S*}(1)]\}, \end{aligned} \tag{C·20d}$$

$$\begin{cases} C_2=C_2(1). \\ A_2=(-p/W)C_2(1), \end{cases}$$

$$C_2(1)=(1/12)[q^2|\mathfrak{Z}(-1)|^2+p^2\lambda_1|\mathfrak{Z}(-2)|^2], \tag{C·21}$$

Definitions:

$$\mathfrak{Z}(-k)=z_0(-k)+2\delta(1+\gamma_1)(\chi_0(k)-\sqrt{2}\,\zeta_0'(k))/\sqrt{15}\,(k+\gamma_k), \tag{A·31}$$

$$\lambda_1=(2+\gamma_2)F_1(Z,W)/2(1+\gamma_1)F_0(Z,W). \tag{3·29}$$

(3) The β-γ directional correlation ($b_{\lambda\lambda'}^{(2)}$):

$$b_{02}{}^{(2)'}=b_{02}{}^{(2)}(\tilde{\xi})+b_{02}{}^{(2)}(1),$$

$$\begin{aligned} b_{02}{}^{(2)}(\tilde{\xi})=&(1/\sqrt{6})\{2(p^2/W)\lambda_2 Re[\mathfrak{Z}(-2)V^*] \\ &+\alpha Z(3/2)p\lambda_3 Im[\mathfrak{Z}(-2)V^*]\}, \end{aligned} \tag{C·22}$$

$$\begin{aligned} b_{02}{}^{(2)}(1)=&(1/\sqrt{6})\{(2/3)p^2 Re[z_0(-2)(w_0(+1)+(3/2)w^T)^*] \\ &-(\pm 2qp^2/3W)Re[z_0(-2)w_0{}^*(-1)] \\ &+\alpha Z(3/8)pW Im[z_0(-2)w^{T*}] \\ &-\alpha Z(3p/8W)Im[z_0(-2)w^{S*}(+1)]\}, \end{aligned} \tag{C·23}$$

$$b_{11}{}^{(2)'}=b_{11}{}^{(2)}(\tilde{\xi})+b_{11}{}^{(2)}(1),$$

$$\begin{aligned} b_{11}{}^{(2)}(\tilde{\xi})=&(-\sqrt{6})\{(2p^2/3W)\lambda_2 Re[\mathfrak{X}(-2)Y^*] \\ &+\alpha Z\tfrac{1}{2}p\lambda_3 Im[\mathfrak{X}(-2)Y^*]\}, \end{aligned} \tag{C·24}$$

$$\begin{aligned} b_{11}{}^{(2)}(1)=&(-\sqrt{6}/9)\{-p^2|x_0(-2)-\tfrac{1}{2}u_0(-2)|^2 \\ &-2p^2 Re[(x_0(1)+u_0(1)+\tfrac{3}{2}u^T)(x_0(-2)-\tfrac{1}{2}u_0(-2))^*] \end{aligned}$$

$$+(\pm 2qp^2/W)\,Re[(x_0(-1)-u_0(-1))(x_0(-2)-\tfrac{1}{2}u_0(-2))^*]$$
$$+\alpha Z(3p^3/2W)\,Im[u_0(1)x_0^*(-2)-\tfrac{1}{2}x_0(1)u_0^*(-2)]$$
$$-\alpha Z(9/4)pW\,Im[(x_0(-2)-\tfrac{1}{2}u_0(-2))u^{T*}]$$
$$+\alpha Z(9p/4W)\,Im[(x_0(-2)-\tfrac{1}{2}u_0(-2))u^{S*}]$$
$$+\alpha Z(\pm 3qp/2)\,Im[u_0(-1)x_0^*(-2)+\tfrac{1}{2}x_0(-1)u_0^*(-2)]\},\quad (C\cdot 25)$$

$$b_{12}^{(2)}=b_{12}^{(2)}(\tilde{\xi})+b_{12}^{(2)}(1)$$
$$b_{12}^{(2)}(\tilde{\xi})=\{-(p^2/W)\lambda_2 Re[\mathfrak{Z}(-2)Y^*]$$
$$-\alpha Z(3p/4)\lambda_3 Im[\mathfrak{Z}(-2)Y^*]\},\quad (C\cdot 26)$$
$$b_{12}^{(2)}(1)=(1/3)\{p^2 Re[(u_0(1)+x_0(1)+\tfrac{3}{2}u^T)z_0^*(-2)]$$
$$-p^2 Re[(x_0(-2)-\tfrac{1}{2}u_0(-2))z_0^*(-2)]$$
$$-(\pm qp^2/W)\,Re[(x_0(-1)-u_0(-1))z_0^*(-2)]$$
$$-\alpha Z(3p^3/4W)\,Im[x_0(+1)z_0^*(-2)]$$
$$+\alpha Z(\pm 3pq/4)\,Im[x_0(-1)z_0^*(-2)]$$
$$+\alpha Z(9pW/4)\,Im[z_0(-2)u^{T*}]$$
$$-\alpha Z(9p/4W)\,Im[z_0(-2)u^{S*}]\},\quad (C\cdot 27)$$

$$b_{22}^{(2)}=b_{22}^{(2)}(1)$$
$$b_{22}^{(2)}(1)=-(1/12)(7/2)^{1/2}p^2\lambda_1|\mathfrak{Z}_0(-2)|^2.\quad (C\cdot 28)$$

Definitions:

$$\lambda_2=(\Lambda/8)[(1+\gamma_1)(3+\gamma_2-\gamma_1)\cos(\theta_1-\theta_2)+\Theta(\gamma_2+1-\gamma_1)\sin(\theta_2-\theta_1)],\quad (C\cdot 29)$$
$$\lambda_3=(\Lambda/6)[(1+\gamma_1)(3+\gamma_2-\gamma_1)\Theta^{-1}\sin(\theta_2-\theta_1)-(\gamma_2+1-\gamma_1)\cos(\theta_2-\theta_1)],\quad (C\cdot 30)$$
$$\Lambda=2\sqrt{F_1}/(1+\gamma_1)\sqrt{F_0},\quad \text{(See (3·29))},\quad (C\cdot 31)$$
$$\Theta=\alpha ZW/p,\quad (C\cdot 32)$$

θ_i is given in (B·24).

(4) **The β-γ correlation measuring circular polarization of γ-ray ($b_{\lambda\lambda'}^{(1)}$ and $b_{\lambda\lambda'}^{(3)}$):**

$$b_{01}^{(1)}=(-p/W)\{b_{01}^{(1)}(\tilde{\xi}^2)+b_{01}^{(1)}(\tilde{\xi})+b_{01}^{(1)}(1)\},$$
$$b_{01}^{(1)}(\tilde{\xi}^2)=2Re[VY^*],\quad (C\cdot 33)$$
$$b_{01}^{(1)}(\tilde{\xi})=(2/3)\{WRe[\mathfrak{W}_{+T}Y^*-\mathfrak{U}_{+T}V^*-10\delta\Gamma_1(\tilde{\xi}')^{-1}VY^*]$$
$$-2W\lambda_4 Re[\mathfrak{X}(-2)V^*]$$
$$-(\pm q)\,Re[\mathfrak{W}_{-}Y^*+\mathfrak{U}_{-}V^*]$$
$$+\alpha Zp^{-1}Im[\mathfrak{W}_{-}Y^*-\mathfrak{U}_{-}V^*]$$

$$-\alpha Z(4+3p^2)\,(2p)^{-1}\lambda_5 Im[\mathfrak{X}(-2)V^*]$$
$$+\alpha Z\tfrac{3}{2}p^{-1}\delta(1+\gamma_1)\xi' Im[r_0^S(-1)\gamma_0^{-1}(-1)+\gamma_0^S(-1)r_0^*(-1)]\},$$
(C·34)

$$b_{11}^{(1)}=(-p/W)\,\{b_{11}^{(1)}(\xi^2)+b_{11}^{(1)}(\xi)+b_{11}^{(1)}(1)\},$$
$$b_{11}^{(1)}(\xi^2)=-\sqrt{2}\,|Y|^2,$$ (C·35)
$$b_{11}^{(1)}(\xi)=(\sqrt{8}/3)\,\{WRe[(\mathfrak{U}_{+T}+5\delta\Gamma_1\xi'^{-1}Y2)\,Y^*]$$
$$-W\lambda_4 Re[\mathfrak{X}(-2)\,Y^*]$$
$$+(\pm q)\,Re[\mathfrak{U}_{-}Y^*]$$
$$+\alpha Z p^{-1}Im[\mathfrak{U}_{-S}Y^*]$$
$$+\alpha Z p^{-1}3\xi(1+\gamma_1)^{-2}Im[u_0^S(+1)x_0^*(+1)+x_0^S(+1)u_0^*(+1)]$$
$$-\alpha Z(4+3p^2)\,(4p)^{-1}\lambda_5 Im[\mathfrak{X}(-2)\,Y^*]\},$$ (C·36)

$$b_{12}^{(1)}=(-p/W)[b_{12}^{(1)}(\xi)+b_{12}^{(1)}(1)],$$
$$b_{12}^{(1)}(\xi)=(-(5/3)^{1/2})\,\{W\lambda_4 Re[\mathfrak{Z}(-2)\,Y^*]$$
$$+\alpha Z(4+3p)\,(4p)^{-1}\lambda_5 Im[\mathfrak{Z}(-2)\,Y^*],$$ (C·37)

$$b_{12}^{(3)}=(-p/W)\,b_{12}^{(3)}(1),$$ (C·38)

$$b_{22}^{(1)}=(-p/W)\,b_{22}^{(1)}(1),$$
$$b_{22}^{(1)}(1)=(\sqrt{10}/24)\,\{q^2|\mathfrak{Z}(-1)|^2+(3/5)p^2\lambda_1|\mathfrak{Z}(-2)|^2\},$$ (C·39)

$$b_{22}^{(3)}=(-p/W)\,b_{22}^{(3)}(1),$$
$$b_{22}^{(3)}(1)=(-1/2\sqrt{10})\,p^2\lambda_1|\mathfrak{Z}(-2)|^2,$$ (C·40)

Definitions:

$$\lambda_4=(\Lambda/8)[(1+\gamma_1)\,(3+\gamma_2-\gamma_1)\cos(\theta_2-\theta_1)$$
$$+(\alpha Z)^2(\gamma_2+1-\gamma_1)\,\Theta^{-1}\sin(\theta_2-\theta_1)],$$ (C·41a)

$$\lambda_5=(\Lambda p^2/(8+6p^2))[(1+\gamma_1)\,(\gamma_2+3-\gamma_1)\,(W/p)^2\Theta^{-1}\sin(\theta_2-\theta_1)$$
$$-(\gamma_2+1-\gamma_1)\cos(\theta_2-\theta_1)].$$ (C·41b)

(5) The β-γ correlation measuring longitudinal polarization of β-ray ($a_{\lambda\lambda'}^{(2)}$):

$$a_{02}^{(2)}=(-p/W)\,\{a_{02}^{(2)}(\xi)+a_{02}^{(2)}(1)\},$$
$$a_{02}^{(2)}(\xi)=(2/3)^{1/2}\{W\lambda_4 Re[\mathfrak{Z}(-2)V^*]$$
$$+\alpha Z(4+3p^2)\,(4p)^{-1}\lambda_5 Im[\mathfrak{Z}(-2)V^*]\},$$ (C·42)

$$a_{11}^{(2)}=(-p/W)\,\{a_{11}^{(2)}(\xi)+a_{11}^{(2)}(1)\},$$
$$a_{11}^{(2)}(\xi)=-(8/3)^{1/2}\{W\lambda_4 Re[\mathfrak{X}(-2)\,Y^*]$$
$$+\alpha Z(4+3p^2)\,(4p)^{-1}\lambda_5 Im[\mathfrak{X}(-2)\,Y^*]\},$$ (C·43)

$$a_{12}^{(2)} = (-p/W)\{a_{12}^{(2)}(\tilde{\xi}) + a_{12}^{(2)}(1)\},$$

$$a_{12}^{(2)}(\tilde{\xi}) = \{-W\lambda_4 Re[\mathfrak{Z}(-2)Y^*] - \alpha Z(4+3p^2)(4p)^{-1}\lambda_5 Im[\mathfrak{Z}(-2)Y^*]\}, \quad (C\cdot 44)$$

$$a_{22}^{(2)} = (-p/W)b_{22}^{(2)}(1). \quad (C\cdot 45)$$

(6) **The transverse β polarization in the plane of β and γ $(d_{\gamma\gamma}^{(2)})$:**

$$d_{02}^{(2)} = d_{02}^{(20)}(\tilde{\xi}) + d_{02}^{(2)}(1)$$

$$d_{02}^{(2)}(\tilde{\xi}) = -(3/2)^{1/2}\{(p/W)\lambda_6 Re[\mathfrak{Z}(-2)V^*] + \alpha Z\lambda_7 Im[\mathfrak{Z}(-2)V^*]\}, \quad (C\cdot 46)$$

$$d_{02}^{(2)}(1) = -(1/6)^{1/2}\{-(\pm qp/W)Re[z_0(-2)w_0^*(-1)] + \tfrac{3}{2}pRe[(w^T - w^S)z_0^*(-2)] - \tfrac{3}{2}\alpha ZW\, Im[w^T z_0^*(-2)] + \tfrac{3}{2}\alpha Z(3p^2+4)(4W)^{-1} Im[w^S z_0^*(-2)]\}, \quad (C\cdot 47)$$

$$d_{11}^{(2)} = d_{11}^{(2)}(\tilde{\xi}) + d_{11}^{(2)}(1)$$

$$d_{11}^{(1)}(\tilde{\xi}) = \sqrt{6}\{(p/W)\lambda_6 Re[\mathfrak{X}(-2)Y^*] + \alpha Z\lambda_7 Im[\mathfrak{X}(-2)Y^*]\}, \quad (C\cdot 48)$$

$$d_{11}^{(2)}(1) = (2/3)^{1/2}\{(\pm qp/W)Re[(x_0(-1) - u_0(-1))(x_0(-2) - \tfrac{1}{2}u_0(-2))^*] - \tfrac{3}{2}pRe[(u^T - u^S)(x_0(-2) - \tfrac{1}{2}u_0(-2))^*] + \tfrac{3}{2}\alpha ZW\, Im[u^T(x_0(-2) - \tfrac{1}{2}u_0(-2))^*] - \tfrac{3}{2}\alpha Z(3p^2+4)(4W)^{-1} Im[u^S(x_0(-2) - \tfrac{1}{2}u_0(-2))^*] - \alpha Zp^2(4W)^{-1} Im[x_0(-2)u_0^*(+1) + \tfrac{1}{2}x_0(+1)u_0^*(-2)] - \alpha Z(\pm q)\, Im[x_0(-2)u_0^*(-1) - \tfrac{1}{2}x_0(-1)u_0^*(-2)]\}, \quad (C\cdot 49)$$

$$d_{12}^{(2)} = d_{12}^{(2)}\tilde{\xi}) + d_{12}^{(2)}(1),$$

$$d_{12}^{(2)}(\tilde{\xi}) = (3/2)\{(p/W)\lambda_6 Re[\mathfrak{Z}(-2)Y^*] + \alpha Z\lambda_7 Im[\mathfrak{Z}(-2)Y^*]\}, \quad (C\cdot 50)$$

$$d_{12}^{(2)}(1) = (3/2)\{(\pm qp/3W)Re[(x_0(-1) - u_0(-1))z_0^*(-2)] - \tfrac{1}{2}pRe[(u^T - u^S)z_0^*(-2)] + \alpha Z(p^2/12W)\, Im[x_0(+1)z_0^*(-2)] - \alpha Z(\pm q/3W)\, Im[x_0(-1)z_0^*(-2)] + \alpha Z\tfrac{1}{2}W\, Im[u^T z_0^*(-2)] - \alpha Z(3p^2+4)(8W)^{-1} Im[u^S z_0^*(-2)]\}, \quad (C\cdot 51)$$

$$d_{22}^{(2)}=0, \tag{C·52}$$

Definitions :

$$\lambda_6=(\Lambda/8)\,(1+\gamma_1)\,(\gamma_2+1+\gamma_1)\cos(\theta_2-\theta_1), \tag{C·53}$$

$$\lambda_7=(\Lambda/8)\,(1+\gamma_1)\,(\gamma_2+1+\gamma_1)\,\theta^{-1}\sin(\theta_2-\theta_1), \tag{C·54}$$

(7) **The transverse polarization of β-ray perpendicular to the plane of β and γ ($e_{\lambda\lambda'}^{(2)}$):**

$$e_{02}^{(2)}=e_{02}^{(2)}(\xi)+e_{02}^{(2)}(1),$$

$$e_{02}^{(2)}(\xi)=(27/32)^{1/2}\{\alpha Z(p/W)\lambda_8 Re[\mathfrak{Z}(-2)V^*]$$
$$+(\alpha Z)^2\lambda_9 Im[\mathfrak{Z}(-2)V^*]\}, \tag{C·55}$$

$$e_{02}^{(2)}=(3/8)^{1/2}\{+(p^2/W)\,Im[z_0(-1)w^{S*}]$$
$$-a Z(\pm qp/2W)\,Re[w_0(-1)z_0^*(-2)]$$
$$+\alpha Z(3p/4)\,Re[(w^T-w^S)z_0^*(-2)]\}, \tag{C·56}$$

$$e_{11}^{(2)}=e_{11}^{(2)}(\xi)+e_{11}^{(2)}(1),$$

$$e_{11}^{(2)}(\xi)=-(27/8)^{1/2}\{\alpha Z(p/W)\lambda_8 Re[\mathfrak{X}(-2)Y^*]$$
$$+(\alpha Z)^2\lambda_9 Im[\mathfrak{X}(-2)Y^*]\}, \tag{C·57}$$

$$e_{11}^{(2)}(1)=-(3/2)^{1/2}\{(2p^2/3W)\,Im[u_0(+1)x_0^*(-2)-\tfrac{1}{2}x_0(+1)u_0^*(-2)]$$
$$-(p^2/W)\,Im[(x_0(-2)-\tfrac{1}{2}u_0(-2))u^{S*}]$$
$$+\alpha Z(\pm qp/2W)\,Re[(x_0(-1)-u_0(-1))(x_0(-2)-\tfrac{1}{2}u_0(-2))^*]$$
$$-\tfrac{3}{4}\alpha Z p\,Re[(u^T-u^S)(x_0(-2)-\tfrac{1}{2}u_0(-2))^*]\}, \tag{C·58}$$

$$e_{12}^{(2)}=e_{12}^{(2)}(\xi)+e_{12}^{(2)}(1)$$

$$e_{12}^{(2)}=(-9/8)\{\alpha Z(p/W)\lambda_8 Re[\mathfrak{Z}(-2)Y^*]$$
$$+(\alpha Z)^2\lambda_9 Im[\mathfrak{Z}(-2)Y^*]\}, \tag{C·59}$$

$$e_{12}^{(2)}(1)=(-3/4)\{(2p^2/3W)\,Im[x_0(+1)z_0^*(-2)]$$
$$-(p^2/W)\,Im[z_0(-2)u^{S*}]$$
$$+\alpha Z(\pm qp/2W)\,Re[(x_0(-1)-u_0(-1))z_0^*(-2)]$$
$$-\tfrac{3}{4}\alpha Z p\,Re[(u^T-u^S)z_0^*(-2)]\}, \tag{C·60}$$

$$e_{22}^{(2)}=0, \tag{C·61}$$

Definitions :

$$\lambda_8=(\Lambda/6)\,(\gamma_1+\gamma_2+3)\cos(\theta_2-\theta_1), \tag{C·62}$$

$$\lambda_9=(\Lambda/6)\,(\gamma_1+\gamma_2+3)\,\theta^{-1}\sin(\theta_2-\theta_1). \tag{C·63}$$

References

1) T. D. Lee and C. N. Yang, Phys. Rev. **104** (1956), 254.
2) Wu, Ambler, Hayward, Hoppes, and Hudson, Phys. Rev. **105** (1957), 1413.
3) Garwin, Lederman, and Weinrich, Phys. Rev. **105** (1957), 1415.
J. I. Friedman and V. L. Telegdi, Phys. Rev. **105** (1957), 1681.
4) R. Arnowitt and G. Feldman, (to be published).
5) M. Gell-Mann Phys. Rev. **111** (1958), 362. E. J. Konopinski and G. E. Uhlenbeck, Phys. Rev. **48**, 7 (1935).
6) Herrmannsfeldt, Maxson, Stahelin, and Allen, Phys. Rev. **107** (1957), 641.
J. S. Allen, Bull. Amer. Phys. Soc. **3** (1958), 52.
7) B. M. Rustad and S. L. Ruby, Post-deadline papper, Amer. Phys. Soc. Meeting, New York (1958)
Note added in proof; Herrmannsfeldt, Burman, Stahelin, Allen and Braid, Phys. Rev. Letter **1** (1958), 61.
8) Goldhaber, Grodzins, and Sunyar, Phys. Rev. **109** (1958), 1015.
9) T. Ahrens and E. Feenberg, Phys. Rev. **86** (1952), 64.
M. E. Rose and R. K. Osborn, Phys. Rev. **93**, (1954), 1315.
M. Ruderman, Phys. Rev. **89** (1953), 1227.
L. N. Zyrianova, Izvest. Akad. Nauk. S.S.S.R. **20** (1956), 1339.
T. Ahrens, Prog. Theor. Phys. **18**, (1957), 331.
10) Lee, Oehme and Yang, Phys. Rev. **106** (1957), 340.
T. D. Lee and C. N. Yang, *Elementary Particles and Weak Interactions*, Brookhaven National Laboratory (1957).
11) M. Morita and R. S. Morita, Phys. Rev. **107** (1957), 1316; **109** (1958), 2408; **110** (1958), 461; **111** (1958), 1130. M. Morita, R.S.Morita and Yamada, Phys. Rev. **111** (1958), 237.
12) R. B. Curtis and R. R. Lewis. Rev. **107** (1957), 1381; A. Z. Dolginov. Nuclear Phys. **5** (1958), 512.
13) Ambler, Hayward, Hoppes and Hudson, Phys. Rev. **108** (1957), 503; **110** (1958), 787.
14) Jackson, Treiman and Wyld, Phys. Rev. **106** (1957), 517.
15) Frauenfelder, Jackson, & Wyld, Phys. Rev. **110** (1958), 451; R. R. Lewis and R. B. Curtis, Phys. Rev. **110** (1958), 919.
16) F. Boehm and A. H. Wapstra, Phys. Rev. **109** (1958), 456.
17) Burgy, Epstein, Krohn, Novey, Raboy, Ringo, and Telegdi, Phys. Rev. **107** (1957), 1731; **110** (1958), 1214.
18) W. Pauli, Nouvo Cimento, **6** (1957), 204.
W. Pauli, Article in "*Niels Bohr and the Development of Physics*", 1955.
19) S. Kahana and D. L. Pursey, Nuovo Cimento **6**, 1469, (1957).
20) G. Luders, Nuovo Cimento **7**, 171 (1958).
T. D. Lee. Proceedings of the Rehovoth Conference in Israel, September 1957, Interscience Publishers, p. 336.
21) See, for example, C. S. Wu, Proceedings of the Rehovoth Conference in Israel, September 1957, Interscience Publishers, p. 346.
22) See, for example, A. R. Edmonds, "*Angular Momentum in Quantum Mechanics*", Princeton University Press (1957).
23) L. C. Biedenharn and M. E. Rose, Rev. Mod. Phys. **25** (1953), 729.
24) Alder, Stech, and Winther, Phys. Rev. **107** (1957), 728.
25) E. J. Konopinski and G. E. Uhlenbeck, Phys. Rev. **60** (1941), 308.
26) R. B. Curtis and R. R. Lewis, Phys. Rev. **107** (1957), 543.
27) Akhiezer and Beretstetsky, "*Quantum Electrodynamics*" (translation, Office of Technical Services, Washington, D. C.) p. 115.
28) M. E. Rose, Phys. Rev. **82** (1951), 389; M. E. Rose and D. K. Holmes, Phys. Rev. **83** (1951), 190.

29) Z. Matumoto and M. Yamada, Prog. Theor. Phys. **19** (1958), 285.
A. Z. Dolginov and I. N. Toptigin, Nuclear Physics **2** (1956), 147.
P. R. Nataf, Journ de Phys. et Radium **17** (1956), 480.
H. Takebe, Prog. Theor, Phys. **12** (1954), 561.
30) R. Steffen, p. 419 of reference 21.
31) H. Mahmoud & E. Konopinski, **88** (1952) 1266.
32) B. Jacobson & E. Henley, Phys. Rev. **108** (1957), 502.
33) M. Yamada, Prog. Theor. Phys. **10** (1953), 252.
E. A. Plassman & L. M. Langer, Phys. Rev. **96** (1954), 1593.
34) R. R. Lewis, Phys. Rev. **108** (1957), 51.
Fujita, Yamada, Matsumoto, and Nakamura, Phys. Rev. **108** (1957), 1104.
35) Geiger, Ewan, Graham, and MacKenzie, Bull. American Phys. Soc. **3** (1958), 51.
36) E. Greuling and M. Meeks, Phys. Rev. **82** (1951), 399.
37) K. Siegbahn, "*Beta and Gamm Ray Spectroscopy*", Interscience (1955), p. 884.
J. P. Davidson, Phys. Rev. **82**, 48 (1951).
38) "*Table of the Gamma Function for Complex Arguments*", United States National Bureau of Standards Applied Mathematics Series ♯34 (1954). "*Tables of Coulomb Wavefunctions*", United States National Bureau of Standards Applied Mathematics Series ♯17 (1952).

Note added in proof. After we sent this paper, we received a manuscript from I. Iben (to be published in Physical Review), in which he discuss the β-γ correlation phenomena.

In order to get some idea about ξ', introduced in (3·8) of this paper, we propose examination of some special light nuclei (Phys. Rev. Letter **1** (1958), 160). The Coulomb correction appearing in the ξ approximation and in the unique forbidden transition is discussed in Physical Review. (Kotani and Ross, to be published.) Cases where the ξ approximation is not good, are also discussed in another paper by Kotani (to be submitted to Physical Review).

Progress of Theoretical Physics, Vol. 20, No. 5, November 1958

Dispersion Relations and Vertex Properties in Perturbation Theory

K. SYMANZIK

Max-Planck-Institut für Physik, Göttingen, Germany

(Received July 17, 1958)

Dispersion relations for nucleon-nucleon scattering with momentum transfer $2'\Delta < \mu$, for meson-nucleon scattering, and related properties of the meson-nucleon vertex, are shown to hold in every order perturbation theory. This is done by extending a method used by Nambu to discuss the perturbation-theoretical parametric representations of Green's functions. In effect arbitrarily complicated Feynman diagrams are reduced to a few simple types.

Y. Nambu has recently published an interesting article[1] in which he sets out to determine the analytic structure of various Green's functions. The means were the parametric integral representations of these functions obtained in perturbation theory. He showed that these representations implied, in the case of the vertex function, the existence of an analytical continuation in the relevant invariant variable, as is assumed and made use of by numerous authors[2]. Independently Schwinger[3] gave, without proof, similar but more detailed representations of special Green's functions.

In this paper we carry Nambu's method somewhat further to the extent that dispersion relations also result, as well in the familiar case of meson-nucleon scattering as in a case where these have not yet been proven rigorously, namely for nucleon-nucleon scattering.

In section I we derive, following Nambu, the perturbation theoretical parametric representation of the contribution to a Green's function from a single Feynman graph and discuss some of its properties. In the following section we develop Nambu's majorization method and, at the same time, modify it in order to be able to "reduce" complicated graphs to simpler ones in a slightly more refined way than he has. This leads in section III to a perturbation-theoretical proof of vertex properties and dispersion relations for nucleon-nucleon and meson-nucleon scattering. The limitations of the present method, the possible meaning of the result, and its relation to Schwinger's proposal[3] are discussed in the final section.

§ I. Parametric representations of Feynman graph contribution

In this section as well as in the rest of this paper we shall not discuss in detail cases of fields with non-zero spin and renormalization. Both have been done by Nambu, and we shall only mention the resulting modifications as we go along. The renormalization of Feynman parameter integrals has earlier been discussed in great detail by Chisholm[4]

From the parametric representation*1) of the propagation function

$$\Delta F(x) = \frac{i}{(2\pi)^4}\int e^{-ikx}\,\frac{dk}{k^2-m^2+i\varepsilon} = \frac{-i}{8\pi^2}\int_0^\infty d\alpha e^{-i\alpha^2/2-im^2/2\alpha}$$

one easily derives for the Fourier transform of the contribution from a Feynman graph without internal vertices[6)*2] to the vacuum expectation value of an N-fold time-ordered operator product the following expression

$$\delta\Big(\sum_j p_j\Big)\cdot G(p_1\cdots p_N) \equiv \int e^{i\sum p_j x_j} F(x_1\cdots x_N)\,dx_1\cdots dx_N =$$

$$= \text{const.}\;\delta\Big(\sum_j p_j\Big)\cdot\int_0^\infty\cdots\int_0^\infty \Pi d\alpha_{ij}\,\frac{\delta(1-\sum\alpha)}{[D_r(\alpha)]^2\Big[-\dfrac{D_r(p,\alpha)}{D_r(\alpha)}-M^2(\alpha)+i\varepsilon\Big]^c} \qquad (1)$$

If the i th and j th vertex are connected by a line carrying mass $m_{ij}=m_{ji}$ and having parameter $\alpha_{ij}=\alpha_{ji}$, the symbols have the meaning:

$$D_r(\alpha) \equiv \begin{vmatrix} \sum\limits_{i\neq 1}^{N}\alpha_{1i} & -\alpha_{12}\cdots & -\alpha_{1N} \\ -\alpha_{21} & \sum\limits_{i\neq 2}^{N}\alpha_{2i}\cdots & -\alpha_{2N} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ -\alpha_{N1} & -\alpha_{N2}\cdots & \sum\limits_{i\neq N}^{N-1}\alpha_{Ni} \end{vmatrix}_r = \sum_{\text{skel.}}\Pi'\alpha \qquad (2a)$$

$$D_r(p_1\alpha) \equiv \begin{vmatrix} 0 & p_1\cdots\cdots & p_N \\ p_1 & \sum\limits_{i\neq 1}^{N}\alpha_{1i}\cdots\cdots & -\alpha_{1N} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ p_N & -\alpha_{N1}\cdots\cdots & \sum\limits_{i\neq N}^{N-1}\alpha_{Ni} \end{vmatrix} = -\sum_{\text{div.}}(\textstyle\sum' p)^2 D_r(\alpha')\,D_r(\alpha'') \qquad (2\cdot b)$$

$$M^2(\alpha) \equiv \sum_{i<j}{}' \frac{m^2_{ij}}{\alpha_{ij}} \qquad (2\cdot c)$$

The index r of the determinants means that an arbitrary line and corresponding column (but in the case of $D_r(p_1\alpha)$ not the first one) is to be omitted. If some inner lines are missing, the respective α_{ij} and m_{ij} are to be left out. Parallel lines give rise to $\sum_k \alpha_{ij}^{(k)}$ in the D_r and to $\sum_k \dfrac{m_{ij}^{(k)2}}{\alpha_{ij}^{(k)}}$ in $M^2(\alpha)$. If there are internal vertices, the corresponding p are to be set equal to zero.

$D_r(\alpha)$ is a sum of products of the α. Namely, a "skeleton" is a *simply* connect-

ed subset of the original lines such that every vertex is on this skeleton. The sum is over the α-products of all skeletons, each counted once. It vanishes if, and only if, all α on at least one cut that separates the graph into two parts vanish simultaneously. $D_r(p, \alpha)$ is the sum $\sum_{\text{div}}'$ corresponding to all possible divisions of the original Feynman graph into two internally connected parts by removing the lines on a cut. $(\sum' p)^2$ is the partial sum squared of the external momenta attached to one of those two parts. $D_r(\alpha')$ and $D_r(\alpha'')$ have the same meaning with respect to the two resulting graphs as $D_r(\alpha)$ has to the original one. Obviously $-D_r(p, \alpha)$ is (exceptional sets of α-values disregarded) a positive definite quadratic form of the p if these were chosen *euclidean* (that is, with positive semidefinite Gram's determinant). Also $D_r(\alpha)$ is, in general, positive. So $-D_r(p, \alpha)/D_r(\alpha) \equiv J(p, \alpha)$ is a positive definite or, exceptionally, semidefinite quadratic form. It is obvious that the α are, up to a common scale factor, the inverse of the usual Feynman parameters.

The integer c is the convergence parameter of the entire graph and is rendered positive by renormalization. Renormalization also has the effect that the integral (1) is also to be done over some mass values which are, however, always larger than the mass values which already appear in the original graph. Consequently, as will be apparent soon, this modification does not affect the later discussion.

If there are particles with spin there appears, as Nambu has shown, in the numerator of the integral (1) a polynomial in the α and p. Since we are interested in the analytic properties of $G(p_1 \cdots p_N)$ as a function of the p and its singularities (besides singularities in infinity, which we shall not discuss) are exclusively given by the zeros of the denominator, we may disregard the mentioned polynomial. Of course, such singularities given by

$$\frac{J(p, \alpha)}{M^2(\alpha)} \equiv Q(p, \alpha) = 1 \tag{3}$$

will sometimes be eliminated by the α-integration, as we shall see later, or they will be "movable" in the sense that they lead to a movable cut in the plane of some variable constructed from the p. Nevertheless, if that variable is fixed so that $M_\alpha Q(p, \alpha) = 1$ it will in general determine a branch point that can not be removed*[3].

A straightforward investigation of $Q(p, \alpha)$ with the help of the expansions (2) shows*[4] that $Q(p, \alpha) = 1$ when a cut exists dividing the original graph into two connected parts in such a way that $(\sum'' p)^2 = (\sum'' m)^2$ where $\sum'' p$ is the partial sum of the momenta on one side and $\sum'' m$ the total mass on the cut. From the work of Eden[5] and Hamilton we know that these points are indeed branch points if the p are real *Minkowski* vectors*[6]. But, as we shall see, in the case of dispersion relations and vertex properties, the interest is in different situations, namely inside "unphysical region" where, as follows from the present discussion, the branch points can be shifted *below* those well-known thresholds*[6]. So we have to resort to different methods to secure $Q(p, \alpha) \neq 1$. The next section is devoted to such a method.

§ II. Extension of Nambu's majorization method

Nambu has shown that the characteristic quotients have the following property: To a complicated graph with quotient $Q(p, \alpha)$ there can be found simpler graphs with, e.g., quotient $Q_0(p, \beta)$ and functions $\beta(\alpha)$ such that

$$0 \leq Q(p, \alpha) \leq Q_0(p, {}_i\beta(\alpha))$$

for all sets α and real *euclidean* sets p. He gives also two correspondences of this type, namely removal of lines and dissolution of closed loops. We shall later need many more correspondences and therefore state, and prove in the appendix, the following.

Theorem: Let $\mathcal{G}$ and $\mathcal{G}'$ be two connected Feynman graphs with vertices $z_i (i=1 \cdots N)$ and $z'_j (j=1 \cdots N')$, respectively. Let m_{ik} and m'_{jl} be the masses carried by the line ik in $\mathcal{G}$ and jl in $\mathcal{G}'$. If then for arbitrary values of the $z_i = z_i'$ of the external vertices and minimizing positions of the internal vertices always holds

$$\sum_{i<k} m_{ik} |z_i - z_k| \leq \sum_{j<l} m_{jl}' |z_j' - z_l'| \tag{4}$$

then

$$Q'(p, \alpha) \leq \operatorname*{Max}_{\beta} Q(p, \beta) \tag{5}$$

holds for all α and real euclidean sets p.

To decide whether (4) is fulfilled in practical cases it is often convenient to use the following.

Lemma: If $\mathcal{G}$ has no internal vertices, and for each line ik in $\mathcal{G}$ there can be chosen in $\mathcal{G}'$ a number of ways leading from $z_i' = z_i$ to $z_k' = z_k$ and a set of corresponding n_{ik} nonnegative massses $(ik)_\nu (\nu = 1 \cdots n_{ik})$ with $\sum_{\nu=1}^{n_{ik}} (ik)_\nu = m_{ik}$ such that for all pairs $jl \sum_{ik \nu \in jl} (ik)_\nu \leq m_{jl}'$ where $\sum_{ik\nu \in jl}'$ means the summation over all those ways which use the line jl in $\mathcal{G}'$ then there exists a set of functions ${}_i\beta(\alpha')$ such that for all α' and euclidean sets p $Q'(p, \alpha') \leq Q(p, {}_i\beta(\alpha'))$ holds.

The lemma is obviously stronger than the theorem, and it is easy to see that the premise of the lemma implies the premise of the theorem. But since in the following we need only the conclusion (5) of the theorem, we omit the very elementary proof of the lemma.

By removing to infinity some of the z_i considered in the theorem, one easily arrives at the following.

Corollary: A necessary condition for the conclusion (5) of the theorem to hold is that the smallest mass on a cut in $\mathcal{G}$ that divides the graph in two or more parts and separates the external vertices in two given non-empty groups is not larger than the mass on any cut with these properties in $\mathcal{G}'$.

In Fig.1 we list a few correspondence[*7)] formulated for meson theory and needed in the next section. (Except n and o they even fulfill the conditions of the lemma.) Only the closed loop d needs an explanation: 2n is the number of vertices. If they were con-

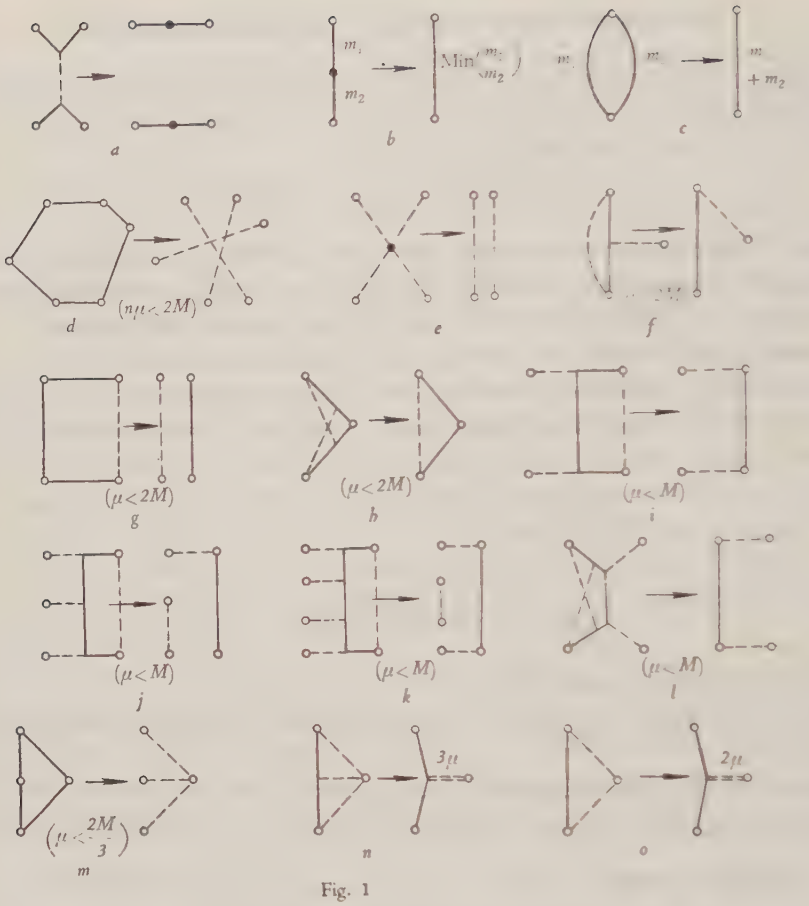


Fig. 1

Fig. 1 A few possible simplifications in the sense of the theorem.

nected in a different order, a correspondingly weaker condition than $n\mu < 2M$ would suffice. Therefore on large loops sufficiently many external meson lines must be eliminated first or the vertices cannot be connected in a simply arbitrary order.

It is clear that, since the condition (4) of the theorem is additive, such a correspondence can be applied in any part of an otherwise arbitrarily complicated graph and, therefor, such a graph can be successively simplified until one achieves some standard forms.

§ III. Vertex properties and dispersion relations

We first consider the vertex function $G(p_1\, p_2\, p_3)$ where $p_1^2 = p_2^2 = M^2$ and the remaining variable is $p_3^2 = k^2$. According to what has been said in section I, $G(k^2)$ will be analytic in the k^2-plane with a cut along the real axis. The cut will go from k^2_{min} to $+\infty$ where k^2_{min} is such that

$$\operatorname{Max}_{\alpha} Q(k^2_{min}, \alpha) = 1. \tag{6}$$

In the interval $0 \leq k^2 \leq 4M^2k^2$ is euclidean (i.e. fulfills the triangle condition) with respect to the other two squares M^2, and the majorization method will suffice to study this region. If we find there a root of (6), and we shall so, then (6) cannot be fulfilled for any k^2 smaller than k^2_{min}, especially not by a negative one, since the left hand side of (6) is an increasing function of k^2_{min}. (A little consideration shows that the case of several vanishing α does not influence this conclusion.) Now in the case of pseudoscalar (or symmetric) mesons every nontrivial vertex can, by the correspondences of Fig. 1, be reduced to the three forms (we assume, of course, $3\mu < 2M$) shown in Figs. 2a-2c. Instead of making an exhausting list of possibilities (it has been done) we only remark that when simplifying graphs one has to pay attention first *not* to eliminate the meson lines that start at the ex-

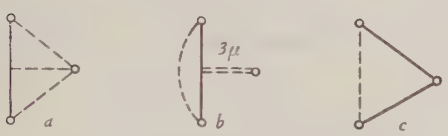


Fig. 2

Fig. 2 Remaining vertex graphs with selection rule.

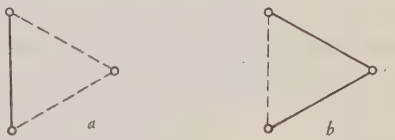


Fig. 3

Fig. 3 Remaining vertex graphs without selection rule.

ternal nucleon vertices[6] and second *not* to diminish thoughtlessly the total mass 3μ which lies between the external meson vertex and the open nucleon line. It is very easy to check that Fig. 2a and 2b give $k^2_{min} = 9\mu^2$ and that 2c gives $k^2_{min} = 4M^2$. In the case of scalar mesons the standard graphs would be those of Fig. 3a-b with thresholds $4\mu^2$ and $4M^2$, respectively.

we now turn to dispersion relations. Since the case of meson-nucleon scattering is rather well covered by exact methods[7] we first study nucleon-nucleon scattering which even turns out to be simpler, with the present method, than meson-nucleon scattering. First we fix the kinematics. When p and q are the initial and $-p'$ and $-q'$ the final four momenta, the spuares of the partial sums that appear in (2b) are :

$$\begin{aligned} &p^2 = p'^2 = q^2 = (p+q+p')^2 = M^2 \\ &(p+p')^2 = -4\Delta^2 \qquad (7\cdot a) \\ &(p+q)^2 = 2M^2 + 2\Delta^2 + 2\omega_V \sqrt{M^2+\Delta^2} \\ &(p'+q)^2 = 2M^2 + 2\Delta^2 - 2\omega_V \sqrt{M^2+\Delta^2}. \end{aligned}$$

The characteristic denominator in (1) takes the form

$$2R(\alpha)\,\omega_V \sqrt{M^2+\Delta^2} + S(\alpha,\Delta^2) - M^2(\alpha) + i\varepsilon. \qquad (8\cdot a)$$

The singularities

$$\omega = \frac{M^2(\alpha) - S(\alpha,\Delta^2)}{2R(\alpha)\sqrt{M^2+\Delta^2}} \mp i\varepsilon \operatorname{sign} R(\alpha)$$

will lie in the ω plane as shown in Fig. 4. In order to have a dispersion relation we

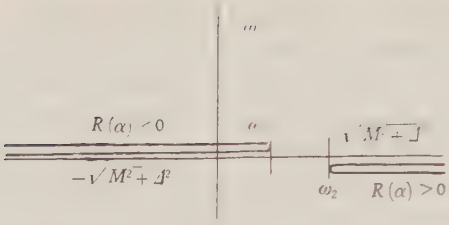


Fig. 4

Fig. 4 Singularities of the scattering amplitude in the ω-plane.

Fig. 5

Fig. 5 Graphs singled out from other nucleon-nucleon scattering graphs.

have to make sure that there is no cut above the real ω axis at $\omega \geq \sqrt{M^2+J^2}$ and none below at $\omega \leq -\sqrt{M^2+J^2}$ in order to have the correct crossing properties. So we have to show that

$$-2|R(\alpha)|(M^2+J^2)+S(\alpha,J^2)<M^2(\alpha).$$

The ideal ways to proceed would be to deduce an estimate for $M^2(\alpha)$ from the method of the previous section. Indeed, (5) would give

$$M'^2(\alpha) \geq \operatorname*{Max}_{p} \operatorname*{Min}_{\beta} \frac{J'(p,\alpha)}{J(p,\beta)} M^2(\beta)$$

where absence of prime indicates a suitably simplified graph. This method would be feasible in the vertex case, where it was not needed, but turns out to be too complicated algebraically in the scattering case. So we have to use the inequality (5) directly without solving for $M^2(\alpha)$. Therefore we replace the set (7·a) by the set

$$\begin{aligned} &p^2=p'^2=q^2=(p+p'+q)^2=M^2+J^2 \\ &(p+p')^2=0 \\ &(p+q)^2=2M^2+2J^2+2\omega_1\sqrt{M^2+J^2} \\ &(p'+q)^2=2M^2+2J^2-2\omega_1\sqrt{M^2+J^2} \end{aligned} \tag{7b}$$

which is euclidean for $|\omega| \leq \omega_1 \sqrt{M^2+J^2}$. Since we have increased the squares, we are sure that if we have no singularity now we did not have one before. Furthermore we can no longer take advantage of the $R(\alpha)$ in (8·a) and have to proceed as follows:

If we can show that

$$2R(\alpha)\omega_1\sqrt{M^2+J^2}+\bar{S}(\alpha,J^2)<M^2(\alpha) \tag{8·b}$$

for all α, where the bar indicates insertion of the above euclidean set, or $\operatorname{Max} Q(\omega,\alpha) < 1$ for ω in some interval

$$-\sqrt{M^2+J^2} \leq \omega_1 < \omega < \omega_2 \leq \sqrt{M^2+J^2},$$

then there is a gap in Fig. 4 from ω_1 to ω_2 and a dispersion relation holds. If this crude method is to work, there must be a gap between the cuts which sets a rather low

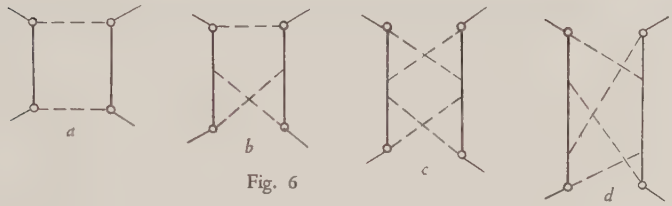


Fig. 6. Remaining nucleon-nucleon scattering graphs.

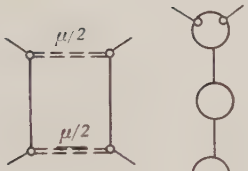


Fig. 7

Fig. 8

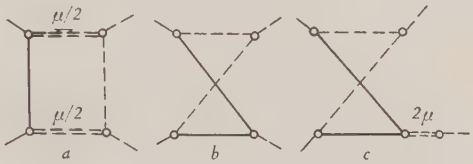


Fig. 9

Fig. 7 Final simplification of the graphs of Fig. 6

Fig. 8 Graphs singled out from other meson-nucleon scattering graphs.

Fig. 9 Remaining meson-nucleon scattering graphs after final simplification.

upper limit to J^2. So it is clear that, e.g., in the well-studied case of equal-particle scattering we cannot obtain as strong results as have already been proven rigorously.

In order to reduce the nucleon-nucleon scattering graphs to simple forms we first separate out the graphs of the general structure, Fig. 5, which are easily analysed with our results for the vertex part and well-known properties or the propagation function. The remaining graphs can be reduced to the graphs Fig. 6a—d and the graphs obtained from them by permuting the external vertices. Instead of listing all possible cases (it has been done) we only remark that the same principle mentioned earlier must be kept in mind here: The meson lines affixed to the external vertices must not be removed, the two open nucleon lines must be preserved as such, and the minimum mass 2μ between them must not be lost.

Since it is still a little bit cumbersome to perform the calculation of ω_1 and ω_2 for the graphs Fig. 6b—d we still reduce all graphs to the single one Fig. 7 and its permuted forms which majorize all of the earlier ones in the sense of the lemma. These final graphs yield for $J^2 < \mu^2/4$ the gap

$$\omega_1 = \frac{2M^2 + 2J^2 - \mu^2}{2\sqrt{M^2 + J^2}} \geq \frac{2M^2 - 2J^2}{2\sqrt{M^2 + J^2}} = \omega_2.$$

ω_2 is just what is expected from the familiar threshold argument whereas in ω_1 there is a replacement of the expected μ by $\mu/2$. This, of course, is due to the unnecessarily crude last step and could presumably be cleared up by the elementary analysis of the graphs Fig. 6b—d.

In the case of meson-nucleon scattering the graphs of the general structure, Fig. 8,

are separated from the others and easily analysed with the help of well-known properties of the vertex function and the nucleon self energy part[*] and obey a dispersion relation for arbitrary J^2. The remaining graphs can be reduced to a few standard types. Without going into details we give for them only the finally simplified forms Fig. 9a–c where lines with artificial masses have already introduced for simplicity. For the same reason, we have not made use of the pseudoscalarity or charge symmetry of the mesons). These graphs are now discussed with the original set of external momenta

$$\begin{aligned} &p^2=p'^2=M^2 \quad k^2=(p-p'-k)^2=\mu^2 \\ &(p+p')^2=-4J^2 \\ &(p+k)^2=M^2+\mu^2-2J^2-2\omega_1\sqrt{M^2+J^2} \\ &(p'+k)^2=M^2+\mu^2+2J^2-2\omega_2\sqrt{M^2+J^2} \end{aligned} \tag{9a}$$

replaced by the set

$$\begin{aligned} &p^2=p'^2=M^2+2J^2 \quad k^2=(p-p'+k)^2=\mu^2 \\ &(p+p')^2=0 \\ &(p+k)^2=M^2+\mu^2+2J^2+2\omega_1\sqrt{M^2+J^2} \\ &(p'+k)^2=M^2+\mu^2-2J^2-2\omega_2\sqrt{M^2+J^2} \end{aligned} \tag{9b}$$

which is euclidean if $|\omega_1|<\mu(M^2-2J^2)^{1/2}(M^2+J^2)^{1/2}$ which case only is of interest here. Since the momentum squared at the external line Fig. 9c is less than $4\mu^2$ the discussion of Fig. 9a and 9b suffices[*10]. The result is: If $J^2<\mu(M+\mu/8$ a dispersion relation holds with a gap from ω_1 to ω_2 where

$$\omega_1=-\frac{M\mu-J^2}{\sqrt{M^2+J^2}} \qquad \omega_2=\frac{M\mu-J^2}{\sqrt{M^2+J^2}} \tag{10}$$

as expected. If

$$\mu(M+\mu/8)-J^2<15M\mu(1-(\mu/8M)^{1/2}/8+\mu^2/64$$

a dispersion relation will still hold but we cannot exclude the possibility that the gap is smaller than that given by (10). For larger values of J^2 we cannot with our method, prove the validity of dispersion relations.

§ IV. Limitations, discussion, and comparison

As has been emphasized, the present method is a very crude one and exhausts in no way the information of the desired character contained on the perturbation formula (1). For instance, the majorization method of section II, when applied to meson-nucleon scattering, certainly does not permit to derive dispersion relations when $J^2>3M\mu(1-(\mu/2M)^{1/2}+\mu^2/2$ because then the cuts in a drawing corresponding to Fig. 4, when calculated from the graph 9c, extend to infinity, that is, we find $\omega_1=+\infty$, $\omega_2=-\infty$. Then a deriv-

ation of dispersion relations requires a discussion of the parametric integrations in (1), and it is the effect of such integrations that is not made use of in the majorization method*[11]. By the same reason, this method does not give any information about the analytic structure of the "imaginary part" whereas a closer analysis of (1) must give such information since Lehmann[7] has already proven, with exact methods, such an analytic property, in the case of the scattering amplitude*[12]. Namely, all results derived *until now* from the general axioms of micro-causality, completeness of certain sets of states, and the asymptotic condition, must also hold in perturbation theory because the properties of Green's functions that have been used there are also properties of single Feynman graphs or, at most, with respect to symmetries, suitable sums of such graphs. So for poor results, methods and their authors are to be blamed but not perturbation theory.

Of course, perturbation theory may give too much analyticity. An example would be the absence of the pole, due to the deuteron, in the perturbation theoretical nucleon-nucleon scattering amplitude which would be proven to be actually present if there were still something similar to a deuteron when $\mu > (\sqrt{2}-1)M$. On the other hand, exact methods will presumably give more information than perturbation theory. As an example we can, unfortunately, only cite the case of a "meson theory" in one space and one time dimension descussed elsewhere[5]. If similar selection rules and spectrum properties are assumed there as in ordinary meson theory, then unitarity urges the renormalized coupling constant g^2 of that paper to be smaller than $32[1-(\mu/2M)^2]^{1/2}M/\mu$*[13].

Nevertheless the result of the present paper as well as Nambu's[1] may stimulate the hope that a more effective use of the completeness condition in the exact methods might enlarge the domain of mass ratios and other parameters for which dispersion relations and similar properties of S-matrix elements can be proven*[14].

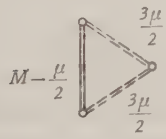


Fig. 10

Fig. 10 A simplification of the vertex graphs Fig. 2.

We finally discuss the relationship of our approach to the general parametric representations of Green's functions proposed by Schwinger.[3]*[15] The theorem of section II and its corollary suggest (but do not prove that it is correct) to replace all graphs by different ones with various masses on the lines that connect the exclusively external vertices such that the cut conditions and only these are fulfilled. In the vertex case, Fig. 2a—c are in the sense of the theorem, reducible to fig. 10. But here one finds that the cut begins already at

$$k^2_{min}=9\mu^2-\theta(2M/5-\mu)\,\mu^2(2M-5\mu)^2/(2M-\mu)^2\approx 8\mu^2$$

which we know is not to be the case in perturbation theory.

Therefore Schwinger's representations*[16]

$$G(p_1\,p_2\,p_3)=\text{const}\cdot\int_{3\mu/2}^{\infty}d\kappa_1^{\,2}\int_{3\mu/2}^{\infty}d\kappa_2^{\,2}\int_{M-\mu/2}^{\infty}d\kappa_3^{\,2}\iiint_0^{\infty}d\alpha_1\,d\alpha_2\,d\alpha_3\,\frac{\delta(1-\alpha_1-\alpha_2-\alpha_3)}{(\alpha_1\alpha_2+\alpha_2\alpha_3+\alpha_3\alpha_1)^2}\,.$$

$$\begin{bmatrix} \alpha_1 p_1^2+\alpha_2 p_2^2+\alpha_3 p_3^2 & -\kappa_1^2 & -\kappa_1^2 & -\kappa_3^2 \\ \alpha_1\alpha_2+\alpha_2\alpha_3+\alpha_3\alpha_1 & \alpha_1 & \alpha_2 & \alpha_3 \end{bmatrix}^{-1} \tag{11}$$

will certainly not be correct unless $\mu(\kappa_1^2,\kappa_2^2,\kappa^2)$ has some special properties that eliminate this prolongated cut.

After the present paper was completed, the author received a preprint by Y. Nambu in which a technique apparently equivalent to that of the present paper was developed and applied to the proof of properties of diverse vertex functions. No derivation of dispersion relations for scattering amplitudes, however, was given.

Appendix

Proof of the theorem of sect. II:

(4) can be written in the following way:

$$\operatorname*{Min}_{\alpha}\operatorname*{Min}_{z_{\text{int}}}\sum_{i<k}\frac{m_{ik}^2}{\alpha_{ik}}\cdot\sum_{i<k}\alpha_{ik}(z_i-z_k)^2\leq$$

$$\leq\operatorname*{Min}_{\alpha'}\operatorname*{Min}_{z'_{\text{int}}}\sum_{j<l}\frac{m_{jl}'^2}{\alpha_{jl}'}\cdot\sum_{j<l}\alpha_{jl}'(z_j'-z_l')^2.$$

Going over to the inverse quadratic forms, it follows

$$\operatorname*{Max}_{\alpha}\frac{J(p,\alpha)}{\sum_{i<k}\frac{m_{ik}^2}{\alpha_{ik}}}\leq\operatorname*{Max}_{\alpha'}\frac{J'(p,\alpha')}{\sum_{j<l}\frac{m_{jl}'^2}{\alpha_{jl}'}}$$

or $Q'(\alpha',p)\leq\operatorname*{Max}_{\alpha}Q(\alpha,p)$ q.e.d.

Here we have made use of the following simple facts:

If a positive definite quadratic form is never smaller than a second quadratic form, then the inverse quadratic form never exceeds the inverse of the second quadratic form.

The inverse form of the minimum of a positive definite quadratic form with respect to some of its variables is equal to the inverse from of the original form with the variables corresponding to the minimizing ones set equal to zero.

Proof: Let the form be

$$\operatorname*{Min}_{\mathcal{Y}}(\mathcal{X}\cdot\mathcal{A}\cdot\mathcal{X}+2\mathcal{X}\cdot\mathcal{B}\cdot\mathcal{Y}+\mathcal{Y}\cdot\mathcal{C}\cdot\mathcal{Y})=\mathcal{X}\cdot(\mathcal{A}-\mathcal{B}\cdot\mathcal{C}^{-1}\cdot\mathcal{B})\cdot\mathcal{X}$$

then the inverse is

$$\mathcal{P}\cdot(\mathcal{A}-\mathcal{B}\cdot\mathcal{C}^{-1}\cdot\mathcal{B}^{\mathsf{T}})^{-1}\mathcal{P}=$$

$$=-|\mathcal{A}-\mathcal{B}\cdot\mathcal{C}^{-1}\cdot\mathcal{B}^{\mathsf{T}}|^{-1}\cdot\begin{vmatrix}0 & \mathcal{P}\\ \mathcal{P} & \mathcal{A}-\mathcal{B}\cdot\mathcal{C}^{-1}\cdot\mathcal{B}^{\mathsf{T}}\end{vmatrix}=$$

$$= - \begin{vmatrix} \mathcal{A} - \mathcal{B} \cdot \mathcal{C}^{-1} \cdot \mathcal{B}^{\top} & 0 \\ \mathcal{B}^{\top} & \mathcal{C} \end{vmatrix}^{-1} \cdot \begin{vmatrix} \not{p} & 0 \\ 0 & \not{p} \end{vmatrix} \cdot \begin{vmatrix} \not{p} & \mathcal{A} - \mathcal{B} \cdot \mathcal{C}^{-1} \cdot \mathcal{B}^{\top} & 0 \\ 0 & \mathcal{B}^{\top} & \mathcal{C} \end{vmatrix} =$$

$$= - \begin{vmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{B}^{\top} & \mathcal{C} \end{vmatrix}^{-1} \cdot \begin{vmatrix} \not{p} & 0 \\ 0 & \not{p} \end{vmatrix} \cdot \begin{vmatrix} \not{p} \mathcal{A} & \not{p} \mathcal{B} \\ \mathcal{B}^{\top} & \mathcal{C} \end{vmatrix}$$

References

1) Y. Nambu, Nuovo Cimento **6** (1957), 1064.
2) E.g. G.F. Chew, R. Karplus, S. Gasiorowicz, F. Zachariasen, Phys. Rev. **110** (1958), 265; V. Glaser, B. Jakišić, Nuovo Cimento **5** (1957), 1197.
3) J. Schwinger, Proceedings of the Seventh Annual Rochester Conference, Sec. IV.
4) J. S. R. Chisholm, Proc. Cambr. Phil. Soc. **48** (1952) 300 a. 518.
5) R. J. Eden, Proc. Roy. Soc. Ld. A **210** (1952), 388.
6) J. Hamilton, Proc. Cambr. Phil. Soc. **48** (1952), 640.
7) E.g. H. Lehmann, Phys. Rev. (in press). This paper contains further references, especially to the fundamental papers by N. N. Bogoliubov et al..
8) K. Symanzik, Nuovo Cimento **5** (1957), 659.

*1) We use the metrik $+1, -1, -1, -1$.

*2) We only consider connected graphs with external lines and, conveniently, self energy parts amputated. Vertices where one or more lines have been amputated are called external ones, all others internal ones. p_i is the sum of the momenta acting on the i th external vertex.

*3) For example if, as in the case of the vertex part, $Q(\kappa^2, \alpha)$ is monotonic, and if in addition $c=1$, then $G(\kappa^2)$ is real on the real κ^2 axis on one side of the point $\kappa_{min}{}^2$ and gets a negative imaginary part when crossing it so that a branch point is unavoidable. (For notation see Sec. III.)

*4) To this end choose all α finite except those on the cut, and let these vanish simultaneously and in proportion to the respective m.

*5) It is easily shown that the criterion for this is that the Gram's determinant has rank less than five and at most one positive eigenvalue.

*6) That a deviation from these "intuitive" thresholds may occur has also been observed by R. Karplus, C. Sommerfield, E. Wichman, by R. Oehme, and by Y. Nambu (preprints).

*7) The correspondences a to d appear already in Nambu's paper, ref. 1.

*8) It is a neglect of this condition which makes Nambu's own proof of threshold properties not quite complete.

*9) These prorerties can, of course, also immediately be deduced by the methods of this paper.

*10) This most plausible result follows from

$$Q(p, \alpha', \alpha) = \frac{\kappa^2 D_r'(\alpha') + \alpha D_r'(p, \alpha')}{\alpha D_r'(\alpha')\,[4\mu^2/\alpha + M'^2(\alpha')]} = \frac{\kappa^2 + \alpha J'(p, \alpha')}{4\mu^2 + \alpha M'^2(\alpha')}$$

where the primed quantities refer to the amputated graph, α is the parameter of the external line, and the formulae 2a-c have been used.

*11) One can show that this integration will indeed have the effect of yielding a dispersion relation in the present case. Let $\mu = M + \varepsilon$ where ε is small. Then the above limit will be $2M^2 + 2M\varepsilon + 0(\varepsilon^2)$. On the other hand, from Lehmann's work (cp. ref.7) follows that we will certainly have a dispersion relation up to $2M^2 + 4M\varepsilon + 0(\varepsilon^2)$.

*12) A certain continuability of the imaginary part of the scattering amplitude as a function of A^2 for fixed total energy follows already from Bogoliubov's work, see ref. 7.

*13) This can be shown as follows; the function

$$D(p) \equiv D(\omega) \frac{p - pB}{p + pB}$$

where $p = (\omega^2 - \mu^2)^{1/2}$ and $pB = i\mu[1 - (\mu^2/2M)^2]^{1/2}$ is regular in the upper half p plane, with $|D(p)| \leq 1$ for real p and

$$D(pB) = -(g^2/32)(\mu/M)[1 - (\mu/2M)^2]^{3/2}.$$

Since $|D(p)|$ does, by assumption, not increase exponentially for $p' \to \infty$, it must obey $|D(p)| \leq 1$ everywhere. The resulting bound for g^2 can, with the help of Poisson-Jensen's formula, even be improved if in some real p intervals lower bounds for the absorptivity $1 - |D(p)|^2$ are known.

*14) A similar opinion is expressed in the papers cited in footnote*6). See also R. Jost, Helv. Phys. Acta **31** (1958), 263.

*15) It has been shown by G. Källen and A. S. Wightman (Dan. Mat. Fys. Medd., in press) that the representations proposed by Schwinger in the addendum to his report at the Seventh Rochester Conference (see ref. 3) violate the causality condition unless the weight function $\rho(\kappa_1^2, \kappa_2^2, \kappa_3^2)$ possesses special properties. This can easily be seen also as follows: The exponent in the integral representation of such an "abnormal" term is

$$-i\lambda_1(x_2 - x_3)^2 - i\lambda_2(x_3 - x_1)^2 - i\lambda_3(x_1 - x_2)^2$$

where e.g. $\lambda_1 = \lambda_3 = 1$, $\lambda_2 = -1/3$ were a permitted triple. If the time coordinates fulfill $x_1^0 = x_2^0 > x_3^0$ then a replacement $x_1^0 \to x_1^0 - i\varepsilon$ produces a *positive* real part whereas it is the spirit of the formula that a non-positive one is produced. The existence of the required analytical continuation can only be secured by special properties of the weight function as stated above. The reason of this circumstance is, of course, that the causality condition is more stringent than Schwinger's "regularity" condition (l.c. equ. 2).

*16) We write (11) with the parameters of the general representation (1).

Progress of Theoretical Physics, Vol. 20, No. 5, November 1958

On the Tests fo Gravitational Theories in Terms of an Artificial Satellite

Hidekazu NARIAI and Yoshio UENO

Research Institute for Theoretical Physics, Hiroshima University, Takehara-machi Hiroshima-ken

(Received July 14, 1958)

We examined whether or not Singer's test and others, if any, would be useful to discriminate between general relativity and other theories of gravitation. Emphasis was laid upon whether curvedness of a space-time could be detectable or not by means of these tests.

§ 1. Introduction

The general theory of relativity is ordinarily regarded as the firmly established classical field theory of gravitation.* According to this theory the metric tensor of a curved space-time is nothing but the gravitational potential in such a way that the motion of a test particle is described by a geodesic of the space-time. Contrary to this idea of Einstein, Gupta[1] proposed another view that Einstein's field equations should be reinterpreted as the non-linear field equations in the flat space-time. Then, has the general theory of relativity been verified by the literally well established experimental or observational facts? Especially is it the case for Einstein's idea of geometrization of gravitation? As is well known, the observational basis of this theory consists simply in the so-called three tests in the solar gravitational field, but phenomena like the gravitational wave have not yet been discovered, although the theory offered a reasonable interpretation of the equivalence** between inertia and gravitation. In addition, the tests in question are not necessarily crucial, apart from the priority of Einstein for the explanation*** and prediction of phenomena. It is because these phenomena can be equally explained in terms not only of the above stated interpretation due to Gupta, but also of different gravitational theories[1] based on the flat space-time. It is very desirable, therefore, that another test is found so that we can predict some result which is different from the one due to general relativity, by using other theories of gravitation permitting us to explain the three tests.

* Except for the standpoint on which gravitation is regarded as a secondary effect of another field[2], say, neutrino or electromagnetic field, we may say that most of the recent attempts to quantize the gravitational field stand on this view-point[3].

** There is a different opinion due to Dicke[3] concerning the limit of the validity of the principle of equivalence.

*** of the anomalous advance amounting to 43″ per century of the periherion of Mercury, for which Newtonian theory was incompetent.

Singer[5] has recently proposed a method of measuring the ratio $\Delta=(dt_{\oplus}/dt_{\infty}-1)$ between the rate of a clock on an artificial satellite and that of a similar clock on the earth, and he has pointed out that this gives the fourth test of the general theory of relativity. According to him, the above Δ is given by the following expression:

$$\Delta=(Gm_{\oplus}/c^2R_{\oplus})\{(0.5-h)/(1+h)\}$$

$$\simeq 6.96\times10^{-10}(0.5-h)(1+h)^{-1}, \qquad (1\cdot1)$$

where h is the height of the artificial satellite in the unit of the radius of the earth $R_{\oplus}$, the other symbols being as usual. Though Singer's test is like that of the red-shift of the spectral lines in the gravitational field of the sun or white dwarfs, there is a considerable difference; *i. e.*, contrary to the latter which is strongly influenced by the Doppler effect due to thermal and turbulent motions of stellar atmospheres, such a complexity of the situation may not arise in the former case.*

Singer's formula (1·1) was, however, derived under the assumption that (1) the gravitational field of the earth can be represented by Schwarzschild's space-time and (2) the orbit of a satellite is circular. Since not only the earth is rotating, but has ellipticity, it is necessary to investigate whether the first of the above assumptions is correct or not. But as Hoffmann[6] and Das[7] showed, the effects due to the rotation and ellipticity of the earth are revealed by small quantities of higher order, and in addition the effects due to them have tendency to cancel each other.

In this paper, therefore, we shall investigate at first the influence of the eccentricity and the phase (specified by the eccentric anomaly), after extending Singer's formula (1·1) to the case of an elliptic orbit. Next we shall examine whether the formula thus obtained is characteristic of the general theory of relativity in the view-point of Einstein himself. For this purpose, we shall attack the same problem from the standpoint of different gravitational theories due to, say, Whitehead and Birkhoff.

Moreover we shall inspect another conceivable test in terms of geodesic deviation instead of geodesic motion. (In the case of an artificial satellite, the equation of geodesic deviation can be interpreted as representing the relative motion of the cap to the main body.) Because this equation includes the curvature tensor $R^{\alpha}_{\beta\gamma\delta}$ explicitly, the value of which tells us directly the curvedness of a space-time.

§ 2. Geodesic equation in Schwarzschild's space-time

As is well known, the line-element of Schwarzchild's space-time is given by:

$$ds^2=c^2(1-2Gm/c^2r)dt^2-dr^2/(1-2Gm/c^2r)-r^2(d\theta^2+\sin^2\theta d\varphi^2). \qquad (2\cdot1)$$

If we choose the coordinate system so that a test-particle moves initially in the plane $\theta=\pi/2$, the geodesic equations in this space-time are

* See, for instance, p. 12 of Singer's paper cited as reference 5) in this paper.

$$d\tau = d\varphi \; hu^2(\varphi), \tag{2·2}$$

$$(du/d\varphi)^2 + u^2 = c^2(\beta^2 - 1)/h^2 + (2Gm/h^2)u + (2Gm/c^2)u^3, \tag{2·3}$$

$$dt/d\tau = \beta(1 - 2Gmu/c^2)^{-1}, \tag{2·4}$$

where $u = 1/r$, $d\tau = ds\ c$, h and β are integration constants. The third term on the right-hand side of the orbital equation (2·3) is the well known term in terms of which the advance of the perihelion of Mercury is explained.

In Singer's problem, however, it is sufficient to take Newtonian approximation for the calculation of the orbit of the artificial satellite, *i.e.*, we can adopt

$$(du/d\varphi)^2 + u^2 = c^2(\beta^2 - 1)/h^2 + (2Gm/h^2)u, \tag{2·3′}$$

in place of (2·3). Then from (2·2) and (2·3′) we can obtain for an elliptic orbit,

$$u = (1 + e\cos\varphi)/a(1 - e^2) \;\; = 1/a(1 - e\cos E), \tag{2·5}$$

$$n\tau = E - e\sin E, \quad \tan(\varphi/2) = \sqrt{(1+e)/(1-e)}\,\tan(E/2), \tag{2·6}$$

$$\beta^2 = 1 - Gm/c^2a \tag{2·7}$$

with usual notations; here we assume $\varphi = 0$ at $\tau = 0$ for convenience.

Though it is sufficient to take the above approximation in the calculation of the orbital motion of the artificial satellite, the essence of Singer's problem lies in considering the deviation of $dt/d\tau$ from unity. Accordingly, inserting (2·5) and (2·7) into the right-hand side of (2·4), and neglecting the terms higher than $(Gm/c^2a)^2$ compared with unity, we obtain

$$dt/d\tau = 1 + \frac{Gm}{c^2a}\left[\frac{2}{1 - e\cos E} - \frac{1}{2}\right]. \tag{2·8}$$

Considering (2·6), we can integrate the above equation as follows:

$$t = \tau + (Gm/c^2a)\,(1/2n)\,(3E + e\sin E). \tag{2·9}$$

§ 3. Generalization of Singer's formula for Δ

As was shown by Hoffmann, effects due to the rotation and ellipticity of the earth can be ignored. So we shall assume that the gravitational field of the earth can be expressed by Schwarzschild's space-time. Then we can take the value of $m_\oplus$ as the value of m appearing in § 2. Accordingly we obtain from (2·8):

$$\frac{dt_s}{d\tau} = 1 + \frac{Gm_\oplus}{c^2R_\oplus}\,\frac{(1+h)^{-1}}{2}\left(\frac{3 + e\cos E}{1 - e\cos E}\right) \tag{3·1}$$

where $a = (1+h)R_\oplus$, $R_\oplus$ = the radius of the earth, t_s is the reading of the clock on the artificial satellite. On the other hand the expression of the reading $t_\oplus$ of the clock which is at rest on the earth can be given, from (2·1) with $m = m_\oplus$, by:

$$dt_{\oplus}/d\tau = 1 + (Gm_{\oplus}/c^2 R_{\oplus}). \tag{3·2}$$

From the above two expressions, we calculate Δ corresponding to (1·1) as:

$$\Delta = (6.96\times 10^{-10})\left[(\psi(e\,;E) - h)/(1+h)\right], \tag{3·3}$$

with

$$\psi(e\,;E) = (1/2)\,(1+3e\cos E)\,(1-e\cos E). \tag{3·4}$$

Especially when $e=0$ (circular orbit) $\psi(0\,;E)=0.5$, so that (3·3) reduces to Singer's formula (1·1).

Now for brevity's sake, we assume that an observer on the earth is situated on the orbital plane of the artificial satellite, *i.e.* on the equatorial plane of the earth. Then, the conditions securing him to observe the artificial satellite are reduced to the one that the perigee distance $a(1-e)$ is larger than the radius of the earth $R_{\oplus}$, *i.e.* h must satisfy the following inequality:

$$h \geq e/(1-e). \tag{3·5}$$

Next we shall study how Δ depends on the eccentricity e and the eccentric anomaly E of the orbit of the artificial satellite. As is easily seen from (3·3), the sign of Δ depends on the magnitude of $\psi(e\,;E)$ compared with h. $\psi(e\,;E)$ is, however, such a periodic function of E that it decreases monotonously in the phase $(0,\ \pi)$ and increases in $(\pi,\ 2\pi)$, if the value of e is fixed. Thus we obtain from (3·4):

$$\begin{aligned}
&\psi_{\max} = \psi(e\,;0) = \frac{1}{2}\left(\frac{1+3e}{1-e}\right),\\
&\psi_{\text{int}} = \psi(e\,;\pi/2) = 0.5, \qquad \text{irrespective of the value of } e,\\
&\psi_{\min} = \psi(e\,;\pi) = \frac{1}{2}\left(\frac{1-3e}{1+e}\right).
\end{aligned} \tag{3·6}$$

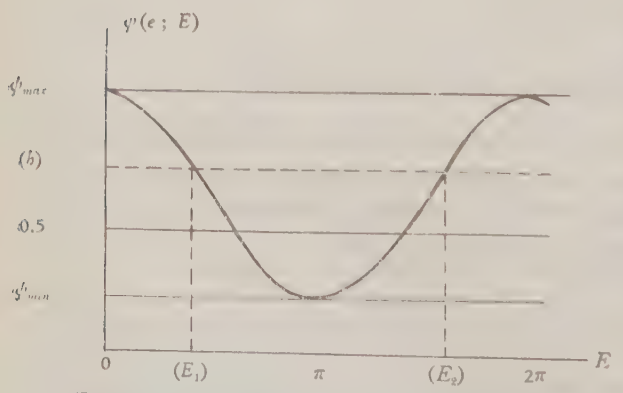


Fig. 1 The relation between ψ and E for the case $e<1/3$. E_1 and E_2 denote the values of E for which $\Delta=0$.

As is easily seen, when $e\neq 0$, $\psi_{\text{int}}=0.5$ is splitted up and down into $\psi_{\max}$ and $\psi_{\min}$ respectively. And the sign of $\psi_{\min}$ depends on the value of e, *i.e.* $\psi_{\min}\gtrless 0$ when $e\lesssim 1/3$.

From the above, signs of Δ can be specified as follows:

$$\begin{aligned}
&\Delta > 0 \text{ (red-shift), when}\\
&h < \psi_{\min} \text{ (if } e < 1/3),\\
&\Delta < 0 \text{ (blue-shift), when}\\
&h > \psi_{\max}.
\end{aligned} \tag{3·7}$$

And when $\zeta'_{\max} > h > \zeta'_{\min}$, the sign of Δ varies as the phase of E. As an example we illustrate such a situation for the case $e < 1/3$ in Fig. 1.

§ 4. Gravitational theories based on the flat space-time due to Whitehead and Birkhoff

In this section we shall deal with Whitehead's theory* and Birkhoff's one* as representative theories which are based on the flat space-time and at the same time by which we can interpret the so-called three tests in the solar gravitational field similarly to the general theory of relativity.

(W-theory). This is an action-at-a-distance theory obtained by reforming Newton's inverse-square law of gravitation in a Lorentz covariant fashion. Therefore this must be automatically rejected when the quantization of the gravitational field is concerned. However, no theory would be better than this, if we could test the appropriateness of the theory without considering the problem of quantization. Meanwhile, Schild[8] showed that we could construct various kinds of action-at-a-distance theories of Whitehead's type by starting from Schwarzschild's line-element. The essential point of his procedure consists in transforming the geodesic equation by the use of quasi-Newtonian coordinate systems and making each term in the equation thus transformed correspond to the Lorentz covariant quantity in the flat space-time. In addition, he asserted that, contrary to general relativity, these theories predict in general the presence of a secular acceleration** of the center of mass in two-body problem.

(B-theory). Contrary to the W-theory, this is a kind of tensorial gravitational field theory. Of this theory, however, there exists Weyl's[9] critique, and further Gupta[10] remarked that difficulty apppears in the quantization problem, *i.e.* the energy of the gravitational field does not take a positive definite value. But it is to be noticed that these critiques are concerned not with pointing out the discrepancy of the theoretical consequences from the observational evidences, but with formalistic points and quantization problem. The original interpretation of the red-shift of spectral lines is, however, *ad hoc* (due to the energy loss of photon). But this was explained by Moshinsky[11] more naturally by taking account of the gravitational effect on the Maxwell and Dirac equations. Moreover, it is to be remarked that the secular acceleration of the center of mass of two bodies takes place also in this theory (just the same form as the one derived erroneously by Levi Civita in general relativity).

In the following we shall study the corresponding formulae of these theories to (3·3) with (3·4) derived by the general theory of relativity.

§ 5. One-body problem in W-theory

Now we shall consider one-body problem which corresponds to Schwarzschild's solution in general relativity. Taking the coordinate system with origin at the gravitating

* In the following we shall designate them as W- & B-theories, respectively.

** It is very difficult, however, to judge whether or not such a secular acceleration exists from the observation of binary stars.

mass, we can describe the Lagrangian of a test particle in the gravitational field from which the equation of motion can be derived:

$$2L=\left(1-\frac{2Gm}{c^2r}\right)\dot{t}^2+\frac{4Gm}{c^3r}\dot{r}\dot{t}-\frac{1}{c^2}\left[\left(1+\frac{2Gm}{c^2r}\right)\dot{r}^2-r^2(\dot{\theta}^2-\sin^2\theta\dot{\varphi}^2)\right], \quad (5\cdot1)$$

where a dot denotes differentiation with respect to parameter τ, and the limitation to τ is specified by the condition $2L=1$.

If we take $\theta=\pi/2$ as in § 2, the Eulerian equation of (5·1) becomes

$$d\tau=d\varphi/hu^2(\varphi), \quad (5\cdot2)$$

$$(du/d\varphi)^2+u^2=c^2(\beta^2-1)/h^2+(2Gm/h^2)u+(2Gm/c^2)u^3, \quad (5\cdot3)$$

$$dt/d\tau=\left\{\beta+\frac{2Gmh}{c^3}u\frac{du}{d\varphi}\right\}\left(1-\frac{2Gm}{c^2}u\right)^{-1}, \quad (5\cdot4)$$

where $u=1/r$, h and β are integration constants the meanings of which are the same as in § 2.

As is easily seen, if we ignore the difference in the physical meaning of τ, (5·2) and (5·3) are of the same form as (2·2) and (2·3), and (5·4) is obtained from (2·4) by means of putting $\{\ \}$ in place of β. Moreover, as in § 2, it is allowed in the Newtonian approximation to put $t=\tau$ in the above equation. Then inserting (2·5)-(2·7) in § 2 into (5·4) and neglecting the terms higher than $(Gm/c^2a)^2$, we obtain:

$$\frac{dt}{d\tau}=1+\frac{Gm}{c^2a}\left[\frac{2}{1-e\cos E}-\frac{1}{2}\right]-(\pm)\left(\frac{Gm}{c^2a}\right)^{\frac{3}{2}}\frac{2e\sin E}{(1-e\cos E)^2}, \quad (5\cdot5)$$

where $(\pm)$ denotes the signs of h. On the other hand, for the test particle which is at rest in the gravitational field, we get from (5·1)

$$dt/d\tau=1+(Gm/c^2R), \quad (5\cdot6)$$

where $r=R=const.$

Applying the above two equations to the problem of the artificial satellite, we obtain the expression of Δ with the same procedure as in § 3 as follows:

$$\Delta=6.96\times10^{10}\left[\frac{\varphi'(e;E)}{1+h}-\frac{h}{1+h}\right]-(\pm)1.54\times10^{-24}(1+h)^{\frac{3}{2}}\frac{2e\sin E}{(1-e\cos E)^2}, \quad (5\cdot7)$$

where $\varphi'(e;E)$ is given by (3·4). Comparing the above formula with (3·3), we find that the second term on the right-hand side is a correction term. But we can neglect this term, because it is very small compared with the first term and, furthermore, it vanishes when $e=0$. Therefore (5·7) and (3·3) are practically equivalent.

Now we shall show that such a similarity of (5·7) to (3·3) is not accidental. Putting $d\tau=ds/c$, we obtain from (5·1)

$$ds^2=c^2\left(1-\frac{2Gm}{c^2r}\right)dt^2+\frac{4Gm}{cr}drdt-\left(1+\frac{2Gm}{c^2r}\right)dr^2-r^2(d\theta^2+\sin^2\theta d\varphi^2). \quad (5\cdot8)$$

Reinterpreting parameter τ as proper time in general relativity, we can regard the above

expression as the line-element of a spherically symmetric space-time. Then, after the following coordinate transformation

$$\bar{t}=t+\frac{2Gm}{c^3}\log(c^2r/2Gm-1), \tag{5·9}$$

(5·8) reduces to

$$ds^2=c^2\left(1-\frac{2Gm}{c^2r}\right)d\bar{t}^2-\frac{dr^2}{1-\frac{2Gm}{c^2r}}-r^2(d\theta^2+\sin^2\theta d\varphi^2), \tag{5·10}$$

which is nothing but Schwarzschild's line-element. (cf. (2·1)). In short the above procedure is reciprocal to Schild's. And, if we interpret (5·7) from the standpoint of general relativity, the appearance of the second term on the right-hand side is due to the fact that the coordinate time t specified by (5·9) is used in place of $\bar{t}$.

§ 6. One-body problem in B-theory

While the one-body problem in W-theory has an intimate connection with Schwarzschild's space-time, the situation is different with the corresponding problem in B-theory: (1) The trajectories of test particles are not geodesics in any curved space-time, *i. e.* the specification of the equations of motion of a test particle is made under another proposal.* (2) Contrary to W-theory, the following relation is an integral of the equation of motion:

$$1=(dt/d\tau)^2-\{(dr/d\tau)^2+r^2(d\theta/d\tau)^2+r^2\sin^2\theta(d\varphi/d\tau)^2\}c^{-2}, \tag{6·1}$$

specifying the flat space-time.

Now, the explicit form of the equations of motion is as follows when $\theta=\pi/2$:

$$dt=\frac{d\varphi}{hu^2(\varphi)}\exp\left(\frac{2Gm}{c^2}u\right), \tag{6·2}$$

$$(du/d\varphi)^2+u^2=(c/h)^2\exp[2Gmu/c^2]\{\exp[2Gmu/c^2]-\beta^{-2}\}, \tag{6·3}$$

$$dt/d\tau=\beta\exp[Gmu/c^2] \tag{6·4}$$

where $u=1/r$, h and β are integration constants.

Comparing (6·2)-(6·4) with (2·2)-(2·4), we can easily understand their differences. Nevertheless, it is to be remarked that the advance of the perihelion of Mercury and the deflection of light ray at solar limb can be derived without *ad hoc* hypothesis.

Then, what is the matter with Singer's test? Inserting the Newtonian approximate

* See the article of Barajas[12]. When we confine ourselves to the one-body problem the equations of motion can be derived from the following variational principle:

$$\delta\int\exp\left(-\frac{2Gm}{c^2r}\right)(c^2+v^2)\,dt=0,$$

which is the same as that proposed by A. G. Walker[13] from the standpoint of Milne's kinematic relativity.

solutions (2·5)-(2·7) of (6·2) and (6·3) into (6·4) and neglecting higher order terms, we obtain

$$\frac{dt}{d\tau}=1+\frac{Gm}{c^2a}\left[\frac{1}{1-e\cos E}-\frac{1}{2}\right]. \tag{6·5}$$

On the other hand, considering that (6·1) is an integral of the equation of motion, we obtain the following expression concerning the particle at rest in gravitational field:

$$dt/d\tau=1\,. \tag{6·6}$$

Comparing the above expressions with (3·1) and (3·2) or (5·5) and (5·6), we can find a remarkable difference. In fact, if we derive the corresponding formula to (3·3) from the above, it follows:

$$\Delta=6.96\times10^{-10}\left[\frac{1+e\cos E}{1-e\cos E}\right]\frac{(1+h)^{-1}}{2}\ , \tag{6·7}$$

which means that Δ is always positive for any e and E. Thus, at first sight Singer's test would be useful to make discrimination between B-theory and general relativity.

It is evidently the case so far as we take the standpoint on which the red-shift of spectral lines is interpreted as caused by the energy loss of photon. Because Singer's test is to directly read the number (counted by a scaler) of ticks of the clock, contrary to the usual measurement of the shift of spectral lines. But it is more appropriate to stand on the view-point due to Moshinsky as for the problem of red-shift (cf. § 5). On this view-point, however, the formula of red-shift can be represented as follows:

$$dt'/dt=1+Gm/c^2r, \tag{6·8}$$

where t' and t are times based on the frequency of spectral line in the presence of and in the absence of the gravitational field, respectively.

The atomic clock is, however, necessary to carry out Singer's test actually. But the time by the atomic clock is based on the unit defined by the reciprocal of frequencies of atomic or molecular lines. Therefore, the relation (6·8) must be taken into account in Singer's problem. Then the following relations must be used instead of (6·5) and (6·6):

$$\frac{dt'}{d\tau}=\frac{dt'}{dt}\,\frac{dt}{d\tau}=1+\frac{Gm}{c^2a}\left[\frac{2}{1-e\cos E}-\frac{1}{2}\right], \tag{6·9}$$

$$\frac{dt'}{d\tau}=1+\frac{Gm}{c^2R}\ , \tag{6·10}$$

where R is constant. Comparing these equations with (3·1) and (3·2), we can find that they are completely the same. Consequently the corresponding formula to (3·3) is identical with (3·3) itself.

Lastly it must be pointed out that the above procedure specified by (6·8) is also applied to Gupta's formalism mentioned before (cf. § 1).

§ 7. Geodesic deviation in Schwarzschild's space-time and its corresponding one in B-theory

As is well known, whether a space-time is flat or not depends not on the values of metric tensor $g_{\mu\nu}$ and its first derivative but only on the fact whether curvature tensor $R^{\mu}_{\alpha\lambda\beta}=0$ or $\neq 0$. For this reason any prediction made by the geodesic equation only does not tell us directly whether the space-time is curved or not, because in this equation $g_{\mu\nu}$ and at most its first derivative are contained. In this respect, Pirani[14] emphasized in his theory of gravitational radiation an importance of the following equation of geodesic deviation :[15]

$$\frac{\partial^2\eta^{\mu}}{\partial s^2}+R^{\mu}_{\alpha\lambda\beta}\lambda^{\alpha}\eta^{\lambda}\lambda^{\beta}=0, \tag{7·1}$$

where ∂ denotes covariant derivative, and λ^{α} is a tangential vector along one geodesic such that $\lambda_{\alpha}\lambda^{\alpha}=1$ and η^{α} is a vector perpendicular to λ^{α} ($\lambda_{\alpha}\eta^{\alpha}=0$).

Then does the equation (7·1) play a significant role also in a static gravitational field such as Schwarzschild's ? Unfortunately it would no be the case, since there exist the following situations : (i) In contrast with the gravitational radiation whose essence consists in the appearance of the discontinuity of $R^{\mu}_{\alpha\lambda\beta}$, there is no such discontinuity in Schwarzshild's space-time, (ii) We can only estimate the mixture of the second term on the left-hand side of (7·1) and $(\partial^2\eta^{\mu}/\partial s^2-d^2\eta^{\mu}/ds^2)$, which is of the same order as the former, *i.e.* we cannot separate the former term from the latter in practice.

Moreover, almost the same equation as (7·1) would be derived also from, say, B-theory, if we take account of the physical meaning of (7·1). For this equation can be interpreted as expressing the relative motion of one test particle on a geodesic line to another one on the neighbouring geodesic line (such a relative motion can be treated even from the standpoint of Newtonian dynamics).

In the following we shall clarify that such a conjecture is valid in the spherically symmetric and static gravitational field, assuming one of geodesic curves as a circular orbit (for the sake of mathematical simplicity).

(General relativity). The solution of (2·2)-(2·4) for a circular orbit is as follows :

$$\frac{d\varphi}{dt}=\left(\frac{Gm}{r_0^3}\right)^{1/2},\quad \frac{dt}{d\tau}=\left(1-\frac{3Gm}{c^2r_0}\right)^{-1}, \tag{7·2}$$

where $r_0=\text{const.}$ Considering (2·1) and (7·2), we obtain the following after calculating (7·1) :

$$\begin{aligned}
&\ddot{X}+n_0^2\left(1-\frac{6Gm}{c^2r_0}\right)\dot{X}=0,\\
&\ddot{Z}+2n_0\left(1-\frac{2Gm}{c^2r_0}\right)^{1/2}\dot{X}=0,\\
&\ddot{Y}+n_0^2Y=0,
\end{aligned} \tag{7·3}$$

where (X, Y, Z) are physical components of (η^1, η^2, η^3), a dot denotes differentiation with respect to t and $n_0=(Gm\ r_0^{-3})^{1/2}$. As is easily seen, if we put parentheses appeared in the first and second equations of (7·3) as unity, then these equations are Newtonian counterparts. And the number 6 in the factor of parentheses of the first equation is familiar in Einstein's formula for the advance of the perihelion of Mercury, *i.e.* $\Delta\bar{\omega}=\dfrac{6\pi Gm}{c^2a(1-e^2)}$ per revolution.

(B-theory). The solution of (6·2)-(6·4) for a circular orbit is:

$$\frac{d\varphi}{dt}=\left(\frac{Gm}{r_0^3}\right)^{1/2}\left(1-\frac{Gm}{c^2r_0}\right)^{-1/2},\qquad \frac{dt}{d\tau}=\left(1-\frac{Gm}{c^2r_0}\right)^{1/2}\Big/\left(1-\frac{2Gm}{c^2r_0}\right)^{1/2}. \tag{7·4}$$

Now we shall calculate the motion of the second test particle Q relative to the first one P referring to the rotating coordinate system (X', Y', Z') as illustrated in Fig. 2. Then, in terms of the method of small oscillation, we obtain

$$\ddot{X}+n^2\left(1-\frac{2Gm}{c^2r_0}\right)\left(1-\frac{4Gm}{c^2r_0}\right)X=0,$$
$$\ddot{Z}+2n\left(1-\frac{Gm}{c^2r_0}\right)\dot{X}=0, \tag{7·5}$$
$$\ddot{Y}+n^2Y=0,$$

where (X, Y, Z) are the same quantities as those in (7·3) and $n=\dot{\varphi}$.

Taking into account that n in (7·5) corresponds to n_0 in (7·3), we can see that these two systems of equations coincide with each other, if we ignore the terms higher than $(Gm/c^2r_0)^2$ compared with 1. Thus it may be said that there appears neither difference between (7·1) and the corresponding equation in B-theory, nor additional prediction from (7·1) to those due to the geodesic equation; the effect due to the factor $\left(1-\dfrac{6Gm}{c^2r_0}\right)$ in the first equation of (7·3) is equivalent to the advance of perihelion in geodesic motions.

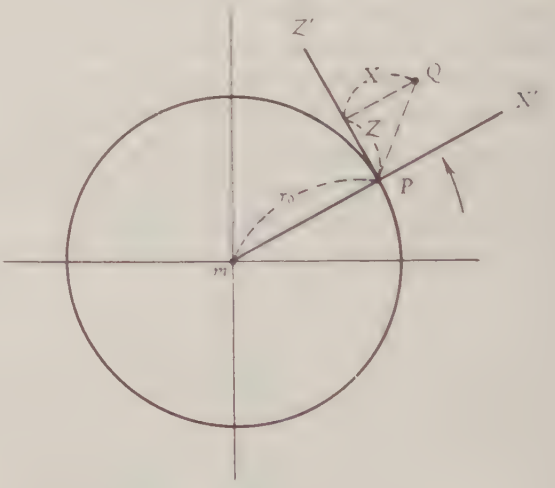


Fig. 2

§ 8. Conclusions

From the above consideration we may conclude that Singer's test is also useless to

discriminate between (1) the orthodox interpretation of gravitation in general relativity and another due to Gupta, and (2) general relativity and W- & B-theories, because the same formula (3·3) can be derived with some proper, but not *ad hoc* procedure in terms of either theory. Moreover, it turns out to be clear that the test due to geodesic deviation is also useless so far as the static gravitational field is concerned.

Then do these circumstances originate from the very nature of the gravitational field? Perhaps they do not. The question whether a space-time is flat or not must have real significance just in the same way that the surface of the earth is closed (concept in-the-large). In this respect, the facts clarified in this paper are simply that it is next to impossible to decide whether a space-time is curved or not so long as the static gravitational field is concerned. Thus the problem is led to finding other objects or methods, if any, suitable to verify whether a space-time is curved or not.

Eventually the investigation of the gravitational field of the celestial bodies may not be suitable for this purpose, because it is not clear whether the properties can be observed or not, if the difference of the properties in-the-large between Schwarzschild's and flat space-time become known.

On the other hand, in cosmology it is evidently not so and the consideration in-the-large would play an important role. It is not desirable, however, to use cosmology as the test of a gravitational theory[16]. If so, except the problem of quantization of the gravitational field, there is little probability of finding out the object by which the appropriateness of the theory is discriminated.

Nevertheless, W-theory is based on an action-at-a-distance and B-theory encounters a difficulty in quantization procedure so long as Gupta's critique is accepted. So there is left only the general theory of relativity. Thus, the only way we could take is either to follow the line of thought due to Gupta et al.[17] or to attack the problem with Wheeler's idea[18] which is, though difficult, more faithful to Einstein's thought. But so long as the applicability of general relativity to cosmology is accepted, it seems to us that the latter way is more plausible. For if we take such a view-point, the world model based on the flat space-time is incompetent to explain the red-shift-magnitude relation of galaxies without *ad hoc* hypothesis other than the concept of cosmic expansion.

The authors wish to express their deep gratitude to Professor Y. Mimura for his helpful discussions and interest in this work.

Reference

1) S. N. Gupta, Proc. Roy. Soc. **A65** (1952), 161, 608; Rev. Mod. Phys. **29** (1957), 334.
2) For example, R. H. Dicke, Rev. Mod. Phys. **29** (1957), 355, 363.
3) See various papers appeared in Rev. Mod. Phys. **29** (1957), No. 3 and *Rep. of Chapel Hill Conference on the Role of Gravitation in Physics.* (1957).
4) a) A. N. Whitehead, *Theory of Relativity*, Cambridge (1922), J. L. Synge, Proc. Roy. Soc. **A211** (1952), 303.
   b) G. D. Birkhoff, Proc. Nat. Acad. Sci. **29** (1943), 231; **30** (1944), 324.
   c) R. H. Dicke, loc. cit.
5) S. F. Singer, Phys. Rev. **104** (1956), 11.
6) B. Hoffmann, Phys. Rev. **106** (1957), 358.
7) A. Das, Prog. Theor. Phys. **18** (1957), 554.
8) A. Schild, Proc. Roy. Soc. **A235** (1956), 202.
9) H. Weyl, Amer. J. Math. **66** (1944), 591.
10) S. N. Gupta, Rev. Mod. Phys. **29** (1957), 334.
11) M. Moshinsky, Phys. Rev. **80** (1950), 514.
12) A. Barajas, Proc. Nat. Acad. Sci. **30** (1944), 54,
13) A. G. Walker, Nature **168** (1951), 961.
14) F. A. E. Pirani, Phys. Rev. **105** (1957), 1089.
15) J. L. Synge and A. Schild, *Tensor Calculus*, Toronto (1949).
16) H. Nariai and Y. Ueno, Pub. Astr. Soc. Japan **9** (1957), 179.
17) For example, T. Kimura, Prog. Theor. Phys. **16** (1956), 157, 555.
18) J. A. Wheeler and C. W. Misner, Ann. Phys. **2** (1957), 525.
   C. W. Misner, Rev. Mod. Phys. **29** (1957), 497.

Progress of Theoretical Physics, Vol. 20. No. 5, November 1958

Equilibrium Properties of Classical Electron Gas in Uniform Positive Ion Distribution

Yoshi H. ICHIKAWA

Deparment of Physics, Tohoku University, Sendai

(Received July 21, 1958)

Thermal equilibrium properties of a classical electron gas are investigated by taking into account the plasma oscillation mode. The free energy is calculated in terms of collective coordinate representation. It is shown explicitly that the Debye-Hückel limiting term of the free energy is due to short range correlation part of the Coulomb interaction, and that the long range correlation effect increases the free energy by 22% of the Debye-Hückel term.

§ 1. Introduction

Recent developments in the researches on the fusion reaction draw our attention to the investigation of the physical properties of a fully ionized gas, so called "*plasma.*"[1),2)] Although the subject appears to belong to applied physics, physics of the plasma is a very interesting fertile field even if one limits his interest to purely academic topics. It is essentially a many body problem of charged particles.

In the past years, many researches have been published concerning dynamical behaviour of the plasma,[3)~9)] yet few works have been done about the thermal equilibrium properties of the classical plasma.[10)] Kihara has proposed to treat the plasma as a strong electrolyte and to apply the Debye-Hückel theory in deriving thermodynamical functions of the plasma.[11)] However, from the series of papers[3)~9)] it is well known that the plasma shows very characteristic dynamical behaviour, namely inside the plasma there occurs an oscillation of density fluctuations called the plasma oscillation. Experimental investigations have confirmed the existence of such a plasma oscillation both in arc discharge tubes[12)] and in metals.[13)] Thus, in a thermal equilibrium state of the plasma, it is quite possible that the plasma oscillation mode shares the partition of energy and plays an important role to establish a thermal equilibrium state. Such a case certainly is out of the applicable range of the Debye-Hückel theory, since in the Debye-Hückel theory the possible existence of the plasma oscillation is not accounted for.

To derive the thermal properties of matter from the microscopic statistical mechanics, it is sufficient to evaluate the partition function

$$Z=h^{-3N}\int\cdots\int\cdots d^N\boldsymbol{p}_i d^N\boldsymbol{x}_i \exp\{-\beta H(\cdots\boldsymbol{p}_i\cdots,\ \cdots\boldsymbol{x}_i\cdots)\} \tag{1}$$

with $\beta=1/\kappa T$, κ: Boltzman's constant, where N is the total number of particles, $H(\cdots \boldsymbol{p}_i \cdots, \cdots \boldsymbol{x}_i \cdots)$ is the Hamiltonian of the system. In ordinary gases, since the inter-particle interactions are of short range forces, the cluster expansion method certainly is a powerful method to evaluate the partition function (1). But the method is completely unapplicable to such a long range force as the Coulomb interaction between charged particles. Concerning this difficulty, Mayer[14] has presented a method to treat the long range Coulomb interaction, and Zubarev[15] has introduced auxiliary variables to treat the long range interaction.

In the following sections, we will discuss a possible way to deal with the thermal equilibrium state of the plasma by taking into account the plasma oscillation mode. In section 2, we will present an effective Hamiltonian for the non-degenerate electron gas in a uniform positive ion background. In section 3, the partition function and the free energy will be evaluated. The last section is devoted to discussions of the result, particularly of our method in comparison with that of Mayer and Zubarev.

§ 2. Effective Hamiltonian

According to the analyses of Bohm and Pines,[4] we will present an effective Hamiltonian for a classical system of the plasma oscillation and electrons in the uniform positive ion background. For a quantum mechanical system, Bohm and Pines[5] have derived an effective Hamiltonian which is analogous to the effective Hamiltonian in classical limit, yet they have made an over-simplification which has been sharply criticized by Brueckner and others.[8] We will come back to this point in the last section.

The Hamiltonian for electrons in the uniform positive ion background is given as

$$H=\sum_i \frac{p_i^2}{2m}+2\pi e^2 \sum_k{}' \frac{1}{k^2}\rho_k \rho_{-k}-2\pi n e^2 \sum_k{}' \frac{1}{k^2}, \tag{2}$$

$$\rho_k=V^{-1/2}\sum_i e^{-i\boldsymbol{k}\boldsymbol{x}_i} \tag{3}$$

where V is the total volume of the system, N is the total number of electrons and $n=N/V$. The density fluctuation ρ_k can be separated into a component q_k, which is due to the effect of the Coulomb interaction, and a component η_k, which remains to be present even if the interaction were absent. Then, one can show that under the random phase approximation the q_k executes a simple harmonic oscillation with an eigen frequency ω_k, which is determined through a dispersion relation

$$1=\frac{4\pi e^2}{m}\frac{1}{V}\sum_i{}' \frac{1}{[\omega-(\boldsymbol{k}\boldsymbol{p}_i/m)]^2}. \tag{4}$$

The dispersion relation (4) is determined to eliminate completely the coupling between the q_k component and the η_k component of the density fluctuation. Explicit expressions of q_k and η_k are given as follows,

$$q_k=V^{-1/2}\sum_i \omega_p^2 [\omega^2-(\boldsymbol{k}\boldsymbol{p}_i/m)^2]e^{-i\boldsymbol{k}\boldsymbol{x}_i}, \tag{5}$$

$$\eta_k = V^{-1/2} \sum_i \{1 - \omega_p^2 / [\omega^2 - (\boldsymbol{k}\boldsymbol{p}_i/m)^2]\} e^{-i\boldsymbol{k}\boldsymbol{x}_i} \tag{6}$$

where $\omega_p^2 = 4\pi e^2 n/m$.

Thus, under the random phase approximation, the exact Hamiltonian (2) is approximated by an effective Hamiltonian

$$H_{R.P.A.} = \sum_i \frac{p_i^2}{2m} + 2\pi e^2 \sum_{k>k_c} \frac{1}{k^2} \rho_k \rho_{-k} + 2\pi e^2 \sum_{k>k_c} \frac{1}{k^2} \eta_k \eta_{-k}$$

$$+ 2\pi e^2 \sum_{k<k_c} \frac{1}{k^2} q_k q_{-k} - 2\pi n e^2 \sum \frac{1}{k^2}. \tag{7}$$

The terms $q_k \eta_{-k}$ and $q_{-k} \eta_k$ are eliminated from the exact Hamiltonian according to the random phase approximation and the dispersion relation (4). This reduction of the exact Hamiltonian is a classical version of the procedure used by Sawada[9] in studies of the correlation energy of a high density degenerate electron gas. Knowing that the collective coordinate q_k executes a simple harmonic oscillation, after introducing a Hamiltonian of the harmonic oscillator system H_f given as eq. (9), we can express the $H_{R.P.A.}$ as a sum of Hamiltonian of individual particle motion H_e and Hamiltonian of the collective motion H_f as follows,

$$H_{R.P.A} = H_e + H_f, \tag{7'}$$

$$H_e = \sum \frac{p_i^2}{2m} + 2\pi e^2 \sum_{k>k_c} \frac{1}{k^2} \rho_k \rho_{-k} - 2\pi n e^2 \sum \frac{1}{k^2} +$$

$$+ 2\pi e^2 \sum_{k<k_c} \frac{1}{k^2} \eta_k \eta_{-k} -$$

$$- \sum_{i,j} \sum_{k<k_c} \frac{m}{I(\omega) k^2} \cdot \frac{(\boldsymbol{k}\boldsymbol{v}_i)}{\omega^2 - (\boldsymbol{k}\cdot\boldsymbol{v}_i)^2} \cdot \frac{(\boldsymbol{k}\boldsymbol{v}_j)}{\omega^2 - (\boldsymbol{k}\cdot\boldsymbol{v}_j)^2} e^{-i\boldsymbol{k}(\boldsymbol{x}_i - \boldsymbol{x}_j)}, \tag{8}$$

$$H_f = \sum_{k<k_c} \left\{ \frac{m}{I(\omega) k^2} \dot{q}_k \dot{q}_{-k} + \frac{m}{I(\omega) k^2} \omega^2 q_k q_{-k} \right\} \tag{9}$$

where

$$I(\omega) = \frac{2}{V} \sum_i \omega^2 / (\omega - (\boldsymbol{k}\boldsymbol{v}_i))^2. \tag{10}$$

The factor $I(\omega)$ insures that the potential energy of the system of plasma oscillators is exactly the corresponding part of the Coulomb interaction $2\pi e^2 \sum q_k q_{-k} k^2$, since $m\omega^2/I(\omega) = 2\pi e^2$. The last term of H_e is the counterpart of the kinetic energy of the plasma oscillators. The critical wave number k_c is introduced as a parameter which separates the effect of the Coulomb interaction into the long range part and the short range part. In the later part of this paper, it will be shown that in the high density limit the k_c can be allowed to become infinite without affecting the result.

The kinetic energy term of H_f (9) is added to give an explicit form of the Hamiltonian of a harmonic oscillator system, after finding that the q_k executes a simple harmonic oscillation with the freqency ω. Hence, the counterpart expressed in terms of the indi-

vidual particle coordinates has to be subtracted from the Hamiltonian of the electron system. In the discussions of the ground state properties of the system, these two terms cancel each other; therefore they will not appear explicitly in the partition function (11). However, one must realize that Bohm and Pines have disregarded the $i \neq j$ terms of H_e (8), while the kinetic energy term of H_r (9) is included as a whole to calculate the correlation energy of the degenerate electron gas.

Concerning the potential energy term, after separating the density fluctuation ρ_k into the q_k and η_k components, the $q_k\ \eta_{-k}$ and $q_{-k}\ \eta_k$ terms of the Coulomb interaction have been eliminated by the dispersion relation (4), thus resulting in the simple harmonic oscillation of the component q_k. Yet, it is very essential to realize that the dispersion relation as well as the random phase approximation has no concern with the $\eta_k\ \eta_{-k}$ term, which has been neglected by Bohm and Pines. Thus, it is clear now that the neglect of $\eta_k\ \eta_{-k}$ terms as well as of the $i \neq j$ terms of the last term of H_e (8), committed by Bohm and Pines[3], is extra-simplification over the random phase approximation and that these inconsistent points are responsible for those criticized by Brueckner and others[5].

Finally, we will add a few words to explain physical meaning of the transformation of the Hamiltonian (2) into the effective Hamiltonian (7), (8) and (9). The effect of the transformation insures "diagonalization" of very complicated correlation effects of the long range Coulomb interaction into the form of the collective oscillation. In other words, this transformation enables us to sum up all contributions of the most-divergent clusters which cannot be handled in the scheme of Mayer's theory of ionic solution[1]. In Mayer's theory only the parts of such most divergent clusters, namely the ring clusters, have been summed up and it has been shown that the result gives the Debye-Hückel limiting term.

§ 3. Calculation of the partition function and the free energy

Using the Hamiltonian (7), (8) and (9), the partition function can be expressed as follows,

$$Z = h^{-3N} \exp\{\sum \alpha(k)\} \prod_k \int dC_k\, dC_{-k} \exp\{-F(k)\, C_k\, C_{-k}\} \cdot$$

$$\cdot \prod_{k<k_c} \int dq_k\, dq_{-k} \exp\{-F(k)\, q_k\, q_{-k}\}\ J(C, q) \tag{11}$$

where

$$\alpha(k) = 2\pi n e^2 \beta / k^2 = (k_d^2 / 2k^2) \tag{12}$$

$$F(k) = \alpha(k)/n\,, \tag{13}$$

$$C_k = \eta_k \qquad \text{for } k < k_c, \tag{14}$$

$$= \rho_k \qquad \text{for } k > k_c.$$

The transfomation function $J(C, q)$ is given as

$$J(C, q) = \int d^N \boldsymbol{p}_i \exp\{-\beta \sum p_i^2 / 2m\} \int d^N \boldsymbol{x}_i \cdot$$

$$\cdot \prod_k \delta\left(C_k - V^{-1/2}\sum_i (1-f_k(\boldsymbol{p}_i))\exp(-i\boldsymbol{k}\boldsymbol{x}_i)\right)\delta(C_{-k}-\cdots)\cdot$$

$$\cdot \prod_k \delta\left(q_k - V^{-1/2}\sum_i f_k(\boldsymbol{p}_i)\exp(-i\boldsymbol{k}\boldsymbol{x}_i)\right)\delta(q_{-k}-\cdots) \tag{15}$$

where

$$f_k(\boldsymbol{p}_i) = \omega_p^2/\left[\omega^2 - (\boldsymbol{k}\boldsymbol{p}_i/m)^2\right] \quad \text{for } k<k_c,$$
$$= 0 \qquad\qquad\qquad\qquad \text{for } k>k_c\,. \tag{16}$$

The identity of eq. (11) with an ordinary phase space integral expression of the partition function for the Hamiltonian (7) is self-evident if one carries out the integrations over the collective coordinates $(C_k,\ q_k)$.

Now, our problem is simply to calculate the transformation function $J(C, q)$. Details of the calculation will be described in Appendix. The result is

$$J \equiv \left(\frac{2\pi m}{\beta}\right)^{3N/2} V^N \left(\frac{i}{2\pi}\right)^2 \prod_{k<k_c} \frac{1}{nf_k^2}\exp\left\{-\frac{Q_k^2}{nf_k^2}\right\}\cdot$$

$$\cdot \prod \frac{1}{n(1-f_k)^2}\exp\left\{-\frac{P_k^2}{n(1-f_k)^2}\right\} \tag{17}$$

where Q_k and P_k are defined by eqs. (A, 4c) and (A, 4d). f_k is defined by eq. (A, 23). Thus, we obtain for the partition function,

$$Z = \left(\frac{2\pi m}{h^2\beta}\right)^{3N/2} V^N \exp\{\sum \alpha(k)\}\cdot$$

$$\cdot \prod \{\alpha(k) f_k^2 + 1\}^{-1} \prod \{\alpha(k)\ (1-f_k)^2 + 1\}^{-1}\,. \tag{18}$$

The free energy is

$$F = F_0 - \kappa T \sum_k \{\alpha(k) - \log(\alpha(k)\ (1-f_k)^2 + 1) -$$

$$-\log(\alpha(k) f_k^2 + 1)\} \tag{19}$$

where F_0 is the free energy for an ideal gas.

Now, to calculate the partition function (18) we need an explicit solution of the dispersion relation (10). Though the question on existence of the root, which corresponds to the plasma oscillation, of the dispersion relation is confronted with mathematical difficulties as it has been discussed by Kampen,[16] at least for the high density limit it has an approximate solution

$$\omega^2 = \omega_p^2\left[1 + 3(k/k_d)^2 + 6(k/k_d)^4\right]. \tag{20}$$

Hence, in the following, we will calculate the partition function at the high density limit. However, one must remember that the result (18) is free from any restriction of the magnitude of the density. Then, corresponding to the approximate solution (20), we have the following expression for the f_k,

$$f_k = \{1 + 2(k/k_d)^2 + 6(k/k_d)^4\}^{-1}. \tag{21}$$

Thus, the free energy due to the Coulomb interaction is obtained as

$$\Delta F = -N\kappa T \frac{V}{4\pi N} \frac{k_d^3}{3_1/\bar{2}} \left[1 + \frac{2\sqrt{2}}{\pi} \Psi(\alpha) \right], \tag{22}$$

$$\Psi(\alpha) = \int_0^\alpha dx \cdot 3x^2 \Big\{ \log(1 + (2x)^{-2}) -$$

$$- \log(1 + (1 - f(x)/2x)^2) - \log(1 + (f(x)/2x)^2) \Big\} \tag{23}$$

where $x = k/k_d$, $\alpha = k_c/k_d$ and $f(x) = (1 + 2x^2 + 6x^4)^{-1}$. The factor $(V/4\pi N)(k_d^2/3_1/\bar{2})$

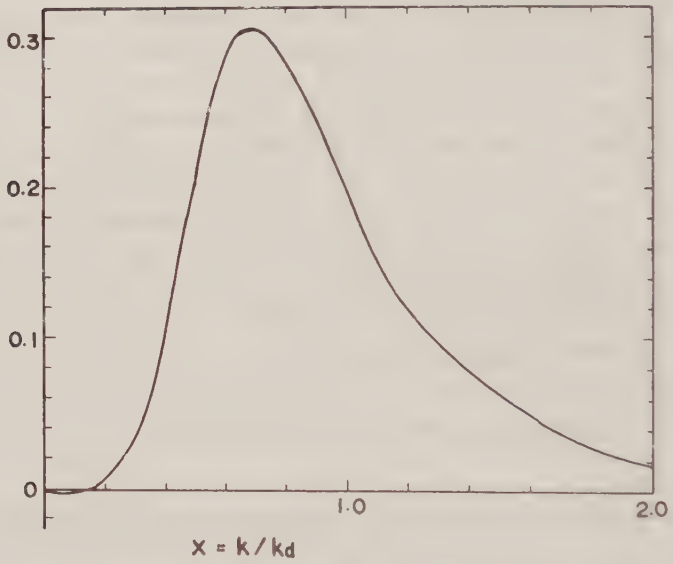


Fig. 1. Variation of the integrand of eq. (23)

is the contribution arising from the upper limit of the k intergration, and it is the Debye-Hückel term. The factor $_1/2$ is a characteristic factor of the present model of the plasma, in which we treat the positive ions as uniform background of positive charge and thus neglect the contribution of the inter-ion interactions. The second term in the bracket of eq. (22) represents the contributions of the long range correlation effect.

It will be instructive to examine the structure of the long range correlation term $\Psi(\alpha)$. The term $\log(1 + (f(x)/2x)^2)$ is the contribution of the plasma oscillation mode and the rest of the integrand of eq. (23) represents the contribution of long range effects of the individual particle correlation. The integrand of $\Psi(\alpha)$ is shown in Fig. 1. The figure shows that the main contribution comes from the region around $k \approx 0.7k_d$. Implications of Fig. 1 are just the same as those of Fig. 2 of reference 8. The integral of $\Psi(\alpha)$ is evaluated numerically and the result is shown in Fig. 2. From the figure, we can see that the effect of the long range correlation is insensitive to the critical wave number k_c, if we take a reasonably large value of k_c. This is in accordance with the

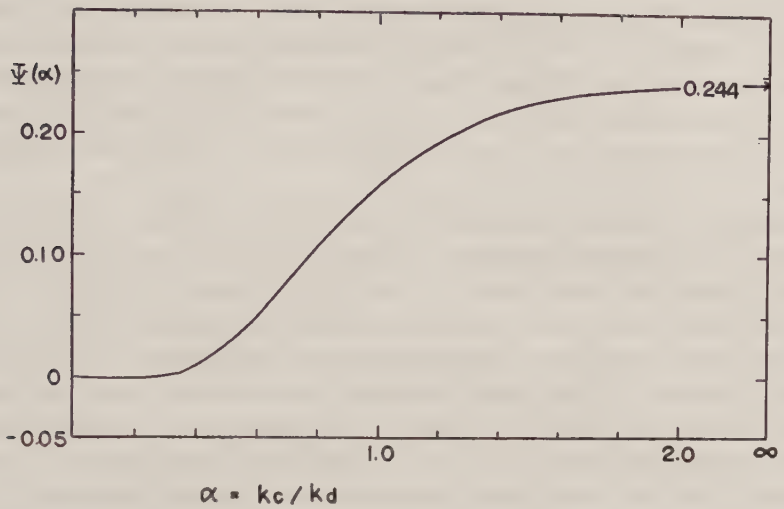


Fig. 2. Variation of the $\Psi(\alpha)$ due to the long range correlation effects.

result obtained by Brueckner and others for the degenerate electron gas at the high density limit.[8] The effect of the long range correlation amounts to be about 22% of the Debye-Hückel limiting term.

§ 4. Discussions

The result obtained in the preceding section shows that the contribution of the long range correlation to the free energy amounts to be about 22% of that of the short range correlation effect. At high temperature, the effect of the Coulomb interaction is so small that the thermal properties of the plasma can be described as an ideal gas. However, effects of the interaction are essential for the transport phenomena. To evaluate the relaxation time, the effective Hamiltonian (7), with (8) and (9), must be used as a basic Hamiltonian, which is transformed from the Hamiltonian (2) after the non-perturbational diagonalization of the strong long range correlation of the Coulomb interaction. Here, one may realize that Zubarev's treatment is a perturbational calculation of the long range correlation in terms of auxiliary variables.

Concerning the effective Hamiltonian (7), with (8) and (9), we will discuss effects of the long range individual particle correlations, namely the effect of the $\gamma_k \gamma_{-k}$ term and of the last term of Hamiltonian (8). Bohm and Pines has disregarded the $\gamma_k \gamma_{-k}$ term by arguing that the term represents part of a screened interaction beyond the screening radius k_c^{-1}. They have also neglected the last term by the same argument except the term with $i=j$, which is involved as modification of the electron mass. However, since the last term is *as a whole* the counterpart of the kinetic energy of the plasma oscillators, it is hard to justify such inconsistent treatment of the last term of Hamiltonian (8). About the effect of the $\gamma_k \gamma_{-k}$ term, the present analysis shows that the presence of the term is essential to get the k_c-insensitive result. The k_c- insensitive result in the high

density limit has been obtained by Brueckner and others[8] by correctly including the term $\hbar\omega_{p'}(e^2 \to 0)$. Hence it is clear now that if one includes the long range individual particle correlation term to the Bohm-Pines effective Hamiltonian, namely if one uses the effective Hamiltonian (7), with (8) and (9) for the degenerate high density electron gas the resulting correlation energy should agree completely with the result obtained by Brueckner and others.[8]

To close the present discussions, a few remarks will be given on the random phase approximation. The random phase approximation picks up the most divergent correlations of particles in which the same amount of momentum is transferred among the particles. On the other hand, Mayer's sum of the ring clusters, which results in the Debye-Hückel term, consists only of the part of the most divergent correlations, or the random-phase-correlations. Therefore, if one gathers the random-phase-correlation parts from other various complicated clusters, they give as a whole the contribution which is also proportional to the square root of the density, $\sqrt{n}$. At the high density limit, according to the present investigation, it has been shown that the contribution is identified with that of the plasma oscillation mode. At the low density limit, the contribution may or may not be identified with that of the plasma oscillation mode, yet so far no theory has carefully checked if the contribution, which is also proportional to $\sqrt{n}$, is negligible compared with that of the ring clusters. We will discuss these points in some details in a separate paper.

It is very essential to extend the present theory to a realistic case, in which the motion of the positive ions is also taken into consideration. It has been shown that in the binary system of electrons and ions there may occur two kinds of collective motion, electronic plasma oscillation and the acoustic oscillation.[7] Detailed studies of these problems will be discussed in another separate paper.

It is a great pleasure to thank Professors T. Kihara, S. Hayakawa and N. Fukuda for their stimulating discussions at the Symposium on Plasma Physics held in May 1958 at Research Institute for Fundamental Physics, Kyoto University. The author is obliged to Professor Hayakawa for his suggestive discussions in the course of preparing this paper. Thanks are also due to Dr. Sumi for his critical discussions.

Appendix

We will present here detail accounts of the derivation of the transformation function $J(C, q)$, eq. (17) and of the partition function, eq. (18). We have an expression (15) for the transformation function $J(C, q)$.

First, let us introduce the Fourier representation of the δ-function as follows,

$$\delta\left(q_k - V^{-1/2}\sum_i f_k(\boldsymbol{p}_i)\exp(-i\boldsymbol{k}\boldsymbol{x}_i)\right)$$

$$= \frac{1}{2\pi}\int d\omega_k \exp\left[i\omega_k\left(q_k - V^{-1/2}\sum_i f_k(\boldsymbol{p}_i)\exp(-i\boldsymbol{k}\boldsymbol{x}_i)\right)\right], \qquad (A\cdot 1)$$

$$\delta\left(C_k - V^{-1/2}\sum_i g_k(\boldsymbol{p}_i)\exp(-i\boldsymbol{k}\boldsymbol{x}_i)\right)$$

$$=\frac{1}{2\pi}\int d\nu_k \exp\left[i\nu_k\left(C_k-V^{-1/2}\sum_i g_k(\boldsymbol{p}_i)\exp(-i\boldsymbol{k}\boldsymbol{x}_i)\right)\right] \tag{A·2}$$

where $g_k(\boldsymbol{p}_i)=1-f_k(\boldsymbol{p}_i)$.

Thus, we have

$$\begin{aligned}
&\delta(C_k-\cdots)\,\delta(C_{-k}-\cdots)\,\delta(q_k-\cdots)\,\delta(q_{-k}-\cdots)\\
&\quad=(2\pi)^{-4}\iiiint d\omega_k\ d\omega_{-k}\ d\nu_k\ d\nu_{-k}\\
&\quad\cdot\exp\left[i(q_k\omega_k+q_{-k}\omega_{-k}+C_k\nu_k+C_{-k}\nu_{-k})\right]\cdot\\
&\quad\cdot\exp\left[-iV^{-1/2}\sum_i\{f_k(\boldsymbol{p}_i)\,(e^{-i\boldsymbol{k}\boldsymbol{x}_i}\omega_k+e^{+i\boldsymbol{k}\boldsymbol{x}_i}\omega_{-k})+\right.\\
&\qquad\left.+g_k(\boldsymbol{p}_i)\,(e^{-i\boldsymbol{k}\boldsymbol{x}_i}\nu_k+e^{+i\boldsymbol{k}\boldsymbol{x}_i}\nu_{-k})\}\right].
\end{aligned} \tag{A·3}$$

Next, we will change the variables $(\omega_k,\ \omega_{-k})$, $(\nu_k,\ \nu_{-k})$, $(q_k,\ q_{-k})$ and $(C_k,\ C_{-k})$ by the following relation,[17] respectively :

$$\left.\begin{aligned}\omega_k+\omega_{-k}&=U_k\sin\varphi_k\\ \omega_k-\omega_{-k}&=+iU_k\cos\phi_k\end{aligned}\right\} \tag{A·4a}$$

$$\left.\begin{aligned}\nu_k+\nu_{-k}&=V_k\sin\phi_k\\ \nu_k-\nu_{-k}&=+iV_k\cos\phi_k\end{aligned}\right\} \tag{A·4b}$$

$$\left.\begin{aligned}q_k+q_{-k}&=2Q_k\cos\sigma_k\\ q_k-q_{-k}&=2iQ_k\sin\sigma_k\end{aligned}\right\} \tag{A·4c}$$

$$\left.\begin{aligned}C_k+C_{-k}&=2P_k\cos\tau_k\\ C_k-C_{-k}&=2iP_k\sin\tau_k\end{aligned}\right\}. \tag{A·4d}$$

The integrations over the original variables are transformed as follows, corresponding to the change of variables (A·4a)–(A·4d),

$$\int d\omega_k\int d\omega_{-k}=(i/2)\int dU_k\ U_k\int d\varphi_k, \tag{A·5a}$$

$$\int d\nu_k\int d\nu_{-k}=(i/2)\int dV_k\ V_k\int d\phi_k, \tag{A·5b}$$

$$\int dq_k\int dq_{-k}=-2i\int dQ_k\ Q_k\int d\sigma_k, \tag{A·5c}$$

$$\int dC_k \int dC_{-k} = -2i \int dP_k \; P_k \int d\tau_k. \tag{A·5d}$$

Then, we get

$$q_k \; \omega_k + q_{-k}\omega_{-k} = Q_k U_k \sin\,(\varphi_k - \sigma_k), \tag{A·6}$$

$$C_k \; \nu_k + C_{-k}\,\nu_{-k} = P_k V_k \sin\,(\phi_k - \tau_k), \tag{A·7}$$

$$\begin{aligned} & f_k(\boldsymbol{p}_i)\,(e^{-i\boldsymbol{k}\boldsymbol{x}_i}\omega_k + e^{+i\boldsymbol{k}\boldsymbol{x}_i}\omega_{-k}) + \\ & \quad + g_k(\boldsymbol{p}_i)\,(e^{-i\boldsymbol{k}\boldsymbol{x}_i}\nu_k + e^{+i\boldsymbol{k}\boldsymbol{x}_i}\nu_{-k}) \\ & = f_k(\boldsymbol{p}_i)\,U_k \sin(\boldsymbol{k}\boldsymbol{x}_i + \varphi_k) + \\ & \quad + g_k(\boldsymbol{p}_i)\,V_k \sin\,(\boldsymbol{k}\boldsymbol{x}_i + \phi_k). \end{aligned} \tag{A·8}$$

By using relations (A.5)–(A.8), the expression (A.3) becomes

$$\begin{aligned} & (2\pi)^{-4}(i/2)^2 \int dU_k \; U_k \int d\varphi_k \int dV_k \; V_k \int d\phi_k \\ & \cdot \exp[iQ_k \; U_k \sin(\varphi_k - \sigma_k)] \exp[iP_k \; V_k \sin\,(\phi_k - \tau_k)]. \\ & \cdot \exp[-iV^{-1/2} \sum_i f_k(p_i)\,U_k \sin(\boldsymbol{k}\boldsymbol{x}_i + \varphi_k)] \\ & \cdot \exp[-iV^{-1/2} \sum g_k(p_i)\,V_k \sin(\boldsymbol{k}\boldsymbol{x}_i + \phi_k)]. \end{aligned} \tag{A·9}$$

Since the factors $\exp\{-F(k)\,C_k\,C_{-k}\}$ and $\exp\{-F(k)\,q_k\,q_{-k}\}$ of the expression of the partition function eq. (11) are indepenent of σ_k and τ_k, as one can see from the changes of variables (A.4c) and (A.4d), which transform these factors into the following,

$$\exp\{-F(k)\,C_k\;C_{-k}\} = \exp\{-F(k)\,Q_k^2\}, \tag{A·10a}$$

$$\exp\{-F(k)\;q_k\;q_{-k}\} = \exp\{-F(k)\,P_k^2\}, \tag{A·10b}$$

we can average (A.9) over the angles σ_k and τ_k. Then, using relations

$$(2\pi)^{-1} \int d\sigma_k \; \exp[iQ_k\,U_k \sin(\varphi_k - \sigma_k)] = J_0(Q_k \; U_k), \tag{A·11a}$$

$$(2\pi)^{-1} \int d\tau_k \exp\,[iP_k\,V_k \sin\,(\phi_k - \tau_k)] = J_0(P_k\,V_k), \tag{A·11b}$$

we can write (A.9) as

$$\begin{aligned} & (2\pi)^{-4}\,(i/2)^2 \int dU_k \;\; U_k \int d\varphi_k \int dV_k\, V_k \int d\phi_k \, J_0(Q_k\;U_k) J_0(P_k\,V_k) \\ & \cdot \exp[-iV^{-1/2} \sum_i f_k(\boldsymbol{p}_i)\,U_k \sin(\boldsymbol{k}\boldsymbol{x}_i + \varphi_k)] \\ & \cdot \exp[-iV^{-1/2} \sum g_k(\boldsymbol{p}_i)\,V_k \sin(\boldsymbol{k}\boldsymbol{x}_i + \phi_k)]. \end{aligned} \tag{A·12}$$

Here, we may carry out the integration over the angles φ_k and ϕ_k approximately as follows,

$$\int d\varphi_k \exp\left[-iV^{-1/2} \sum_i f_k(\boldsymbol{p}_i) U_k \sin(\boldsymbol{k}\boldsymbol{x}_i + \phi_k) \right]$$

$$\simeq 2\pi \left[1 - (U_k^2/4V) \sum_{i,j} f_k(\boldsymbol{p}_i) f_k(\boldsymbol{p}_j) \cos(\boldsymbol{k}\cdot\boldsymbol{x}_i - \boldsymbol{x}_j) \right]$$

$$\simeq 2\pi \exp\left[-(U_k^2/4V) \sum_{i,j} f_k(\boldsymbol{p}_i) f_k(\boldsymbol{p}_j) \cos(\boldsymbol{k}\cdot\boldsymbol{x}_i - \boldsymbol{x}_j) \right], \qquad (A\cdot 13a)$$

$$\int d\phi_k \exp\left[-iV^{-1/2} \sum_i g_k(\boldsymbol{p}_i) V_k \sin(\boldsymbol{k}\boldsymbol{x}_i + \phi_k) \right]$$

$$\simeq 2\pi \exp\left[-(V_k^2/4V) \sum_{i,j} g_k(\boldsymbol{p}_i) g_k(\boldsymbol{p}_j) \cos(\boldsymbol{k}\cdot\boldsymbol{x}_i - \boldsymbol{x}_j) \right]. \qquad (A\cdot 13b)$$

Thus, (A.12) becomes

$$(i/4\pi)^2 \int dU_k\, U_k J_0(Q_k\, U_k) \int dV_k\, V_k J_0(P_k\, V_k).$$

$$\cdot \exp\left[-(4V)^{-1} \sum_{i,j} (U_k^2 f_k(\boldsymbol{p}_i) f_k(\boldsymbol{p}_j) + V_k^2 g_k(\boldsymbol{p}_i) g_k(\boldsymbol{p}_j)) \cos(\boldsymbol{k}\cdot\boldsymbol{x}_i - \boldsymbol{x}_j) \right]. \qquad (A\cdot 14)$$

Now, let us turn back to the expression of $J(C, p)$ and carry out the integrations over the spatial and momentum coordinates. We have

$$J(C, q) \equiv (i/4\pi)^2 \Pi_k \int dU_k\, U_k J_0(Q_k\, U_k) \int dV_k\, V_k J_0(P_k\, V_k).$$

$$\cdot \int d^N \boldsymbol{p}_i \exp\{-\beta \textstyle\sum p_i^2/2m\} \int d^N \boldsymbol{x}_i \cdot$$

$$\cdot \exp\left\{ -(4V)^{-1} \sum_{i,j} (U_k^2 f_k(\boldsymbol{p}_i) f_k(\boldsymbol{p}_j) + V_k^2 g_k(\boldsymbol{p}_i) g_k(\boldsymbol{p}_j)) \cos(\boldsymbol{k}\cdot\boldsymbol{x}_i - \boldsymbol{x}_j) \right\}. \qquad (A\cdot 15)$$

Here the contribution of terms with $i \neq j$ is smaller by a factor $1/\sqrt{N}$ than that of terms with $i=j$, hence we get

$$\left[V \int d\boldsymbol{p} \exp\{-\beta p^2/2m\} \exp\{-(4V)^{-1} (U_k^2 f_k^2(\boldsymbol{p}) + V_k^2 g_k^2(\boldsymbol{p}))\} \right]^N \qquad (A\cdot 16)$$

for the integrand of (A.15). The approximations employed in (A.13) and in (A.16) are exact in calculating the free energy up to the order of $\sqrt{n}$.

A factor of (A.16),

$$\int d\boldsymbol{p} \exp\{-\beta p^2/2m\} \exp\left\{ -(4V)^{-1} (U_k^2 f_k^2(\boldsymbol{p}) + V_k^2 g_k^2(\boldsymbol{p})) \right\}, \qquad (A\cdot 17)$$

can be replaced by

$$(2\pi m/\beta)^{3/2}\exp\left\{-(4V)^{-1}(U_k^2\langle f_k^2\rangle+V_k^2\langle g_k^2\rangle)\right\} \tag{A·18}$$

where $\langle f_k^2\rangle$ and $\langle g_k^2\rangle$ are averages over the momentum space with the weight of exp $\{-\beta p^2/2m\}$. Thus, putting (A.18) into (A.15), we get for $J(C, q)$ the following,

$$\begin{aligned} J(C, q) &= (2\pi m/\beta)^{3N/2}V^N(i/4\pi)^2 \\ &\cdot \prod_k \int dU_k U_k J_0(Q_k U_k)\exp\{-(n/4)\langle f_k^2\rangle U_k^2\} \\ &\cdot \prod_k \int dV_k V_k J_0(P_k V_k)\exp\{-(n/4)\langle g_k^2\rangle V_k^2\}\,. \end{aligned} \tag{A·19}$$

The integrations over U_k and V_k give the final result for $J(C, q)$ as

$$\begin{aligned} J(C, q) &= (2\pi m/\beta)^{3N/2}V^N(i/2\pi)^2 \\ &\cdot \prod_k (n\langle f_k^2\rangle)^{-1}\exp\{-Q_k^2/n\langle f_k^2\rangle\} \\ &\cdot \prod_k (n\langle g_k^2\rangle)^{-1}\exp\{-P_k^2/n\langle g_k^2\rangle\}\,. \end{aligned} \tag{A·20}$$

For the partition function, we get

$$\begin{aligned} Z &= (2\pi m/h^2\beta)^{3N/2}V^N\exp\{\textstyle\sum\alpha(k)\} \\ &\cdot \prod_k \int dQ_k(2Q_k/n\langle f_k^2\rangle)\exp\{-Q_k^2/n\langle f_k^2\rangle\} \\ &\cdot \prod_k \int dP_k(2P_k/n\langle g_k^2\rangle)\exp\{-P_k^2/n\langle g_k^2\rangle\}\,. \end{aligned} \tag{A·21}$$

For the mean values $\langle f_k^2\rangle$ and $\langle g_k^2\rangle$, we will approximate them by expressions

$$\begin{aligned} \langle f_k(p)^2\rangle &= f_k^2, \\ \langle g_k(p)^2\rangle &= (1-f_k)^2 \end{aligned} \tag{A·22}$$

where f_k is defined by

$$f_k=\omega_p^2/[\omega^2-\langle(\boldsymbol{kx}/m)^2\rangle_{\mathrm{av}}]. \tag{A·23}$$

References

1) R. F. Post, Rev. Mod. Phys. **23** (1956), 338.
2) L. Spitzer, Nature **181** (1958), 221.
3) A. A. Vlasov, J. Phys. USSR **9** (1945), 25.
4) D. Bohm and D. Pines, Phys. Rev. **85** (1952), 338.
5) D. Pines and D. Bohm, Phys. Rev. **92** (1953), 609. From an analysis of the kinematical behaviour of a many particle system, Tomonaga laid foundations of the general theory of the collective motion, including those of Bohm and Pines; S. Tomonaga, Prog. Theor. Phys. **13** (1955), 467, 482.
6) D. Pines, Phys. Rev. **92** (1953), 626.
7) V. P. Silin, J. Exper. Theor. Theor. Phys. USSR **23** (1952), 649.
8) K. A. Brueckner, K. Sawada, N. Fukuda and R. Brout, Phys. Rev. **108** (1957), 507.
9) K. Sawada, Phys. Rev. **106** (1957), 372.
10) S. F. Edwards, Phil. Mag. **3** (1958), 119.
11) T. Kihara, Private communication
12) S. Kojima, K. Kato and S. Hagiwara, J. Phys. Soc. Japan **12** (1957), 1277.
13) H. Watanabe, J. Phys. Soc. Japan **11** (1956), 112.
14) J. E. Mayer, J. Chem. Phys. **18** (1950), 1426.
15) D. N. Zubarev, DAN USSR **95** (1954), 757.
16) N. G. Kampen, Physica **23** (1957), 641.
17) A similar method of evaluation has been introduced independently by; J. K. Percus and G. J. Yevick, Phys. Rev. **110** (1958), 1.

Progress of Theoretical Physics, Vol. 20, No. 5, November 1958

Bose-Einstein Lattice Gas Theory of Ferromagnetism

Tohru MORITA

Physics Department, Tokyo Institute of Technology, Tokyo

(Received June 30, 1958)

Several calculations of the temperature dependence of the spontaneous magnetization of ferromagnet which have been published are not in coincidence with each other. In this paper, it is calculated by applying the cluster development of a gas to the spin-deviation gas—the Bose-Einstein lattice gas equivalent to the Heisenberg model of ferromagnetism. The results same as Dyson has given are obtained without meeting with the problems of non-diagonality, kinematical interaction or non-Hermitic Hamiltonian.

§ 1. Introduction

In the previous paper,[1] the author has shown that the problem of the Heisenberg model of ferromagnetism can be reduced to that of a non-ideal Bose-Einstein lattice gas (spin-deviation gas), and that the results of the spin wave theory of Bloch[2] are obtained in the ideal Bose-Einstein lattice gas approximation in our formalism, namely by neglecting the interactions of spin-deviations. The next step will be, of course, to include the effects of binary collisions of spin-deviations. It is performed in this paper.

The main task of this paper is to calculate the effects of the collisions of two spin-deviations, following the line due to Koster and Slater[3] and Van Kranendonk[4] with cares on the attractive interaction of two spin-deviations when they are on nearest neighbor positions: the neglect of this term and of the effect of finite $\boldsymbol{K}$ values has led Van Kranendonk[4] to erroneous results for the case of $S=1/2$. Though Dyson[5] has attributed the erroneous results of Van Kranendonk to the formalism to use the complete set $\prod_{j=1}^{L} a_j^{+n_j}|0\rangle$ which was introduced by Holstein and Primakoff[6] and took up the non-orthogonal complete set $\prod_{j=1}^{L} S_j^{+n_j}|0\rangle$, it is not the case. We will show that Dyson's results are obtained in the formalism of spin-deviation gas theory which is a generalization of the formalism of Van Kranendonk[4] for $S=1/2$ to the case of $S\geq 1$.

§ 2. Bose-Einstein lattice gas equivalent to the Heisenberg model of ferromagnetism

The Hamiltonian of the Heisenberg model of ferromagnet with spin S is as follows:

$$\mathcal{H}_S=-2J\sum_{<jk>}\boldsymbol{S}_j\cdot\boldsymbol{S}_k-g\mu_B H\sum_{j=1}^{L}S_{jz} \tag{2·1}$$

where J is the exchange integral, g the g-factor, μ_B the Bohr magneton, H the magnetic field along z-direction and L the number of lattice sites in our system. The first sum is

over $Lc/2$ nearest neighbor lattice sites pairs, c being the coordination number. This $\mathcal{H}_S$ has been shown to be equivalent to the Hamiltonian of a Bose-Einstein lattice gas:

$$\mathcal{H}_S = -L_C JS^2 - Lg\mu_B H + g\mu_B H \sum_{j=1}^{L} a_j^* a_j + \mathcal{H}_G \tag{2·2}$$

$$\mathcal{H}_G = 2SJ \sum_{<jk>} (a_j^* - a_k^*)(a_j - a_k) - 2J \sum_{<jk>} a_j^* a_j a_k^* a_k + v_0 \sum_{j=1}^{L} a_j^{*2S+1} a_j^{2S+1}$$

$$-2SJ \sum_{<jk>} \left\{ \left[a_j^* \sum_{n=1}^{2S-1} \frac{c_n}{n!} (a_j^{*n} a_j^n + a_k^{*n} a_k^n) a_k + a_j^* \sum_{n=1}^{2S-1} \frac{c_n}{n!} a_j^{*n} a_j^n \sum_{m=1}^{2S-1} \frac{c_m}{m!} a_k^{*m} a_k^m a_k \right] + c.c. \right\} \tag{2·3}$$

with $v_0 = \infty$, where

$$c_1 = -1 + (1 - 1/2S)^{1/2} \tag{2·3′}$$

$$c_2 = 1 - 2(1 - 1/2S)^{1/2} + (1 - 1/S)^{1/2} \tag{2·3″}$$

.....

Here, $\mathcal{H}_G$ commutes with the number of spin-deviations $N = \sum_{j=1}^{L} a_j^* a_j$, and so the free energy F_S of our spin system at temperature $kT = 1/\beta$ in a magnetic field H is related to the grand partition function $\mathcal{Z}_{gr}$ of spin-deviation gas of chemical potential

$$\mu = kT \ln z = -g\mu_B H \tag{2·4}$$

and temperature kT. That is

$$e^{-\beta F_S} = \mathrm{Tr} \exp[-\beta \mathcal{H}_S] = e^{\beta(LcJS^2 + Lg\mu_B H)} \cdot \mathcal{Z}_{gr}. \tag{2·5}$$

Here $\mathcal{Z}_{gr}$ is given by

$$\mathcal{Z}_{gr} = \sum_{N=0}^{\infty} \mathcal{Z}_N z^N \tag{2·6}$$

where $\mathcal{Z}_N$ is the partition function of N spin-deviations:

$$\mathcal{Z}_N = \mathrm{Tr}^{(N)} \exp[-\beta \mathcal{H}_G]. \tag{2·7}$$

The first term of $\mathcal{H}_G$ is the Hamiltonian of an ideal Bose-Einstein lattice gas

$$\mathcal{H}_{\text{ideal } G} = -Ja^2 \sum_{i=1}^{N} \frac{\Delta^2}{\Delta \boldsymbol{r}_i^2} \tag{2·8}$$

written in the formalism of the second quantization,[7] where

$$\frac{\Delta^2}{\Delta \boldsymbol{r}_i^2} \phi(\cdots, \boldsymbol{r}_i, \cdots) = \frac{1}{2} \sum_{s=1}^{c} \frac{1}{\boldsymbol{a}_s^2} [\phi(\cdots, \boldsymbol{r}_i + \boldsymbol{a}_s, \cdots) - 2\phi(\cdots, \boldsymbol{r}_i \cdots) + \phi(\cdots, \boldsymbol{r}_i - \boldsymbol{a}_s, \cdots)], \tag{2·9}$$

$\boldsymbol{a}_c$'s being vectors from a lattice site to its nearest neighbors and $a=|\boldsymbol{a}_c|$. The other terms represent the collisions of spin-deviations.

The magnetization M is related to the number of spin-deviations N by

$$M/M_\infty=1-N/LS \tag{2·10}$$

where $M_\infty=\gamma\hbar\mu_B LS\,V$ is the saturation magnetization, V being the volume of our system. With the introduction of the Weiss molecular field constant $q=2cJ\gamma^2\hbar^2\mu_B^2$, the magnetic field is reduced to

$$H/qM_\infty=-T^\dagger/3\cdot\ln z \tag{2·11}$$

where

$$T^\dagger=3kT/2cSJ. \tag{2·12}$$

In the following the quantities with a dagger are always measured in the unit of $2cSJ/3$. The susceptibility is given by

$$q\chi=q\frac{\partial M}{\partial H}=\frac{3}{LST^\dagger}\frac{\partial N}{\partial\ln z}. \tag{2·13}$$

§ 3. Cluster development

The problem of calculating the spontaneous magnetization has reduced to that of obtaining N/L for $z=1$. Then we must write N/L in terms of z. For this purpose, we need a fugacity development of the grand free energy—logarithm of the grand partition function—, applicable to the Bose-condensation of a quantum gas. At the present, we have one such development due to Dyson.[5] However, it seems somewhat cumbersome. We will here present an alternative, simpler method of fugacity development, which is based on the method of the Ursell development.

We construct the Slater sums in the momentum representation:

$$W_N(\boldsymbol{k}_1,\cdots,\boldsymbol{k}_N)=W_N(\{n_k\})=\frac{1}{\Pi_k n_k!}\sum_P\langle\boldsymbol{k}_1^P,\cdots,\boldsymbol{k}_N^P|\exp[-\beta\mathcal{H}_N]|\boldsymbol{k}_1,\cdots,\boldsymbol{k}_N\rangle \tag{3·1}$$

where

$$|\boldsymbol{k}_1,\cdots,\boldsymbol{k}_N\rangle=|\boldsymbol{k}_1\rangle\cdots|\boldsymbol{k}_N\rangle,$$

$$|\boldsymbol{k}\rangle=V^{-1/2}e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\qquad\text{for a continuous space,}$$

$$|\boldsymbol{k}\rangle=L^{-1/2}e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\qquad\text{for a lattice space.}$$

Now, we apply the Ursell development to $W_N(\boldsymbol{k}_1,\cdots,\boldsymbol{k}_N)$ as follows:

$$W_1(\boldsymbol{k})=S_1(\boldsymbol{k})$$

$$W_2(\boldsymbol{k}_1,\boldsymbol{k}_2)=S_1(\boldsymbol{k}_1)S_1(\boldsymbol{k}_2)+S_2(\boldsymbol{k}_1,\boldsymbol{k}_2)$$

$$W_2(\boldsymbol{k},\boldsymbol{k})=S_1(\boldsymbol{k})^2+1/2!\cdot S_2(\boldsymbol{k},\boldsymbol{k})$$

$$W_3(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)=S_1(\boldsymbol{k}_1)S_1(\boldsymbol{k}_2)S_1(\boldsymbol{k}_3)+S_1(\boldsymbol{k}_1)S_2(\boldsymbol{k}_2,\boldsymbol{k}_3)+S_1(\boldsymbol{k}_2)S_2(\boldsymbol{k}_3,\boldsymbol{k}_1)$$
$$+S_1(\boldsymbol{k}_3)S_2(\boldsymbol{k}_1,\boldsymbol{k}_2)+S_3(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)$$

$$W_3(\boldsymbol{k}, \boldsymbol{k}, \boldsymbol{k}') = S_1(\boldsymbol{k})^2 S_1(\boldsymbol{k}') + S_1(\boldsymbol{k}') \cdot 1/2! \cdot S_2(\boldsymbol{k}, \boldsymbol{k}) + 2S_1(\boldsymbol{k}) S_2(\boldsymbol{k}, \boldsymbol{k}') + \\ + 1/2! \cdot S_3(\boldsymbol{k}, \boldsymbol{k}, \boldsymbol{k}')$$

$$W_3(\boldsymbol{k}, \boldsymbol{k}, \boldsymbol{k}) = S_1(\boldsymbol{k})^3 + 3S_1(\boldsymbol{k}) \cdot 1/2! \cdot S_2(\boldsymbol{k}, \boldsymbol{k}) + 1/3! \cdot S_3(\boldsymbol{k}, \boldsymbol{k}, \boldsymbol{k})$$

.....

where we have assumed that $\boldsymbol{k}_1 \neq \boldsymbol{k}_2, \boldsymbol{k}_3$, $\boldsymbol{k}_2 \neq \boldsymbol{k}_3$ and $\boldsymbol{k} \neq \boldsymbol{k}'$. In general,

$$W_N(\boldsymbol{k}_1, \cdots, \boldsymbol{k}_N) = \sum_{(\sum k m_k = N)} \prod_{i=1}^{m_1} S_1(\boldsymbol{k}_{1i}) \frac{1}{\prod n_k'!} \prod_{k \geq 2} \prod_{j=1}^{m_k} S_k(\boldsymbol{k}_{kj,1}, \cdots, \boldsymbol{k}_{kj,k}) \qquad (3 \cdot 2)^*$$

where n_k' is the total number of $\boldsymbol{k}$ appearing in the arguments of S_k's with $k > 1$ in the respective product of summand. We transform the right-hand side of (3·2) as

$$W_N(\{n_k\}) = \prod_k S_1(\boldsymbol{k})^{n_k} + \sum_{k_1, k_2} \frac{1}{2!} \frac{n_{k_1}(n_{k_2} - \delta_{k_1, k_2})}{(1 + \delta_{k_1, k_2})!} S_2(\boldsymbol{k}_1, \boldsymbol{k}_2) \prod_k S_1(\boldsymbol{k})^{n_k - \delta_{k, k_1} - \delta_{k, k_2}}$$

$$+ \cdots\cdots$$

$$= \Bigg\{ 1 + \sum_{k_1, k_2}{}' \frac{1}{2!} S_2(\boldsymbol{k}_1, \boldsymbol{k}_2) \frac{1}{(1 + \delta_{k_1, k_2})!} \frac{\partial^2}{\partial S_1(\boldsymbol{k}_1) \partial S_1(\boldsymbol{k}_2)}$$

$$+ \sum_{k_1, k_2, k_3} \frac{1}{3!} S_3(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) \frac{1}{\prod n_k'!} \frac{\partial^3}{\partial S_1(\boldsymbol{k}_1) \partial S_1(\boldsymbol{k}_2) \partial S_1(\boldsymbol{k}_3)}$$

$$+ \frac{1}{2!} \sum_{k_1, k_2, k_3, k_4} \frac{1}{2!} S_2(\boldsymbol{k}_1, \boldsymbol{k}_2) \frac{1}{2!} S_2(\boldsymbol{k}_3, \boldsymbol{k}_4) \frac{1}{\prod n_k'!} \frac{\partial^4}{\partial S_1(\boldsymbol{k}_1) \cdots \partial S_1(\boldsymbol{k}_4)}$$

$$+ \sum_{k_1, k_2, k_3, k_4} \frac{1}{4!} S_4(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \boldsymbol{k}_4) \frac{1}{\prod n_k'!} \frac{\partial^4}{\partial S_1(\boldsymbol{k}_1) \cdots \partial S_1(\boldsymbol{k}_4)} + \cdots \Bigg\} \prod_k S_1(\boldsymbol{k})^{n_k}. \qquad (3 \cdot 3)$$

By the aid of the definition of the partition function

$$Z = \sum_{\substack{\{n_k\} \\ \sum n_k = N}} W_N(n_k) , \qquad (3 \cdot 4)$$

the grand partition function is obtained to be

$$Z_{gr} = e^{\sum_{n=2}^{\infty} \sum_{k_1, \cdots, k_n} (1/n!) S_n(\boldsymbol{k}_1, \cdots, \boldsymbol{k}_n) z^n / \prod_{i=1}^n (1 - z S_1(\boldsymbol{k}_i))} \cdot \prod_k \frac{1}{1 - z S_1(\boldsymbol{k})} . \qquad (3 \cdot 5)$$

As the result, we have

$$\ln Z_{gr} = -\sum_k \ln(1 - z S_1(\boldsymbol{k})) + \sum_{n=2}^{\infty} \sum_{k_1, \cdots, k_n} \frac{z^n}{n!} \frac{S_n(\boldsymbol{k}_1, \cdots, \boldsymbol{k}_n)}{\prod_{i=1}^n (1 - z S_1(\boldsymbol{k}_i))} . \qquad (3 \cdot 6)$$

* To be understood as being identical with the Ursell development[8] except for the appearance of the factor $1/\prod n_k'!$.

The number of particles is given by

$$N=z\frac{\partial}{\partial z}\ln Z_{gr}=\sum_{k}\frac{zS_1(\boldsymbol{k})}{1-zS_1(\boldsymbol{k})}$$
$$+\sum_{n=2}^{\infty}\sum_{\boldsymbol{k}_1,\cdots,\boldsymbol{k}_n}\frac{z^n}{n!}\frac{S_n(\boldsymbol{k}_1,\cdots,\boldsymbol{k}_n)}{\prod_{i=1}^{n}(1-zS_1(\boldsymbol{k}_i))}\left(n+\sum_{i=1}^{n}\frac{zS_1(\boldsymbol{k}_i)}{1-zS_1(\boldsymbol{k}_i)}\right) \tag{3·7}$$

As is naturally expected from our Ursell expansion scheme, the first term represents the term for the ideal Bose-Einstein gas and the next sum represents the interactions of particles. In the development by Dyson $S_n(\boldsymbol{k}_1,\cdots,\boldsymbol{k}_n)$ is further divided into a sum of various terms. The correction due to interactions of our lattice gas will be considered here only up to $S_2(\boldsymbol{k}_1,\boldsymbol{k}_2)$.

Before we calculate $S_2(\boldsymbol{k}_1,\boldsymbol{k}_2)$, we will calculate the number of particles $N_S^{(0)}/L$ at $z=1$ for our ideal Bose-Einstein lattice gas, of which one particle spectrum is given by

$$E_k^{(0)}=J\sum_{s=1}^{c}(1-\cos\boldsymbol{k}\cdot\boldsymbol{a}_s) \tag{3·8}$$

where $\boldsymbol{k}=2\pi\boldsymbol{n}/Ld$, $\boldsymbol{n}=(n_x,n_y,n_z)$, n_x,n_y,n_z : integers, is taken to be in the first Brillouin zone. The Slater sum of one particle is given by

$$W_1(\boldsymbol{k})=S_1(\boldsymbol{k})=e^{-\beta E_k^{(0)}}=e^{-\beta J\sum_{s=1}^{c}(1-\cos\boldsymbol{k}\cdot\boldsymbol{a}_s)} \tag{3·9}$$

The number of particles is given by

$$N=\sum_{k}\frac{zS_1(\boldsymbol{k})}{1-zS_1(\boldsymbol{k})}\ , \tag{3·10}$$

from which one gets for a simple cubic lattice

$$N_S^{(0)}/L=2.612\,(T^{\dagger}/2\pi)^{3/2}(1+0.1925T^{\dagger}+0.1112T^{\dagger 2}+\cdots), \tag{3·11}$$

by putting $z=1$. Analogous expressions have been given by Dyson[5] also for f.c.c. and b.c.c. lattices. The first term of (3·11) and those of corresponding formulae for f.c.c. and b.c.c. lattices represent Bloch's $T^{3/2}$ law. In order to obtain the corrections due to the binary collision, we will turn to the problem of two spin-deviations in the next section. The interaction of more than two spin-deviations will not be discussed in this paper and will be retained as a subject of future investigations.

§ 4. Two-body problem

The purpose of this section is to calculate

$$\frac{1}{1+\delta_{k_1,k_2}}S_2(\boldsymbol{k}_1,\boldsymbol{k}_2)=W_2(\boldsymbol{k}_1,\boldsymbol{k}_2)-W_1(\boldsymbol{k}_1)W_1(\boldsymbol{k}_2)\ . \tag{4·1}$$

Eingen-functions of two spin-deviations can be written as

$$\Psi=\sum_{r_1=1}^{L}\sum_{r_2=1}^{L}\psi(\boldsymbol{r}_1,\boldsymbol{r}_2)\,a_{r_1}{}^{*}a_{r_2}{}^{*}\Psi_0\,, \tag{4·2}$$

where Ψ_0 is the state where no spin-deviation is present. The eigenvalue equation for $\psi(\boldsymbol{r}_1,\boldsymbol{r}_2)$ is obtained by operating $\langle\Psi_0 a_{r_1}^{*}a_{r_2}^{*}$ from the left on

$$\mathcal{H}_G\Psi=E\Psi. \tag{4·3}$$

Then we have

$$\mathcal{H}_2\psi(\boldsymbol{r}_1,\boldsymbol{r}_2)=E\psi(\boldsymbol{r}_1,\boldsymbol{r}_2) \tag{4·4}$$

with

$$\mathcal{H}_2\psi(\boldsymbol{r}_1,\boldsymbol{r}_2)=-J\sum_{s=1}^{c}[\psi(\boldsymbol{r}_1+\boldsymbol{a}_s,\boldsymbol{r}_2)-2\psi(\boldsymbol{r}_1,\boldsymbol{r}_2)+\psi(\boldsymbol{r}_1,\boldsymbol{r}_2+\boldsymbol{a}_s)]$$

$$-2J\sum_{s=1}^{c}\delta_{r_1,r_2+a_s}\psi(\boldsymbol{r}_1,\boldsymbol{r}_2)+v_0\delta_{r_1,r_2}\psi(\boldsymbol{r}_1,\boldsymbol{r}_2)\,,\quad v_0=\infty \tag{4·5a}$$

for $S=1/2$, and

$$\mathcal{H}_2\psi(\boldsymbol{r}_1,\boldsymbol{r}_2)=-2SJ\sum_{s=1}^{c}[\psi(\boldsymbol{r}_1+\boldsymbol{a}_s,\boldsymbol{r}_2)-2\psi(\boldsymbol{r}_1,\boldsymbol{r}_2)+\psi(\boldsymbol{r}_1,\boldsymbol{r}_2+\boldsymbol{a}_s)]$$

$$-2J\sum_{s=1}^{c}\delta_{r_1,r_2+a_s}\psi(\boldsymbol{r}_1,\boldsymbol{r}_2)$$

$$-2SJc_1\Big\{\delta_{r_1,r_2}\sum_{s=1}^{c}[\psi(\boldsymbol{r}_1,\boldsymbol{r}_1+\boldsymbol{a}_s)+\psi(\boldsymbol{r}_1+\boldsymbol{a}_s,\boldsymbol{r}_1)]+2\sum_{s=1}^{c}\delta_{r_1,r_2+a_s}\psi(\boldsymbol{r}_1,\boldsymbol{r}_1)\Big\} \tag{4·5b}$$

for $S\geq 1$, where $\delta_{A,B}$ is Kronecker's symbol. Because of the translational invariance of $\mathcal{H}_2$, the solutions have the form

$$\psi(\boldsymbol{r}_1,\boldsymbol{r}_2)=e^{i\boldsymbol{K}\cdot\boldsymbol{R}}\psi(\boldsymbol{r}) \tag{4·6}$$

where

$$\boldsymbol{R}=(\boldsymbol{r}_1+\boldsymbol{r}_2)/2,\ \boldsymbol{r}=\boldsymbol{r}_1-\boldsymbol{r}_2\,. \tag{4·7}$$

The equation for $\psi(\boldsymbol{r})$ for a value of total momentum $\boldsymbol{K}$ is obtained as

$$E\psi(\boldsymbol{r})=-J\sum_{s=1}^{c}\Big\{\cos\frac{\boldsymbol{K}\cdot\boldsymbol{a}_s}{2}[\psi(\boldsymbol{r}+\boldsymbol{a}_s)+\psi(\boldsymbol{r}-\boldsymbol{a}_s)]-2\psi(\boldsymbol{r})\Big\}$$

$$-2J\sum_{s=1}^{c}\delta_{r,a_s}\psi(\boldsymbol{a}_s)+v_0\delta_{r,0}\psi^{(0)}\,,\qquad v_0=\infty \tag{4·8a}$$

for $S=1/2$, and

$$E\psi(\boldsymbol{r})=-2SJ\sum_{s=1}^{c}\Big\{\cos\frac{\boldsymbol{K}\cdot\boldsymbol{a}_s}{2}[\psi(\boldsymbol{r}+\boldsymbol{a}_s)+\psi(\boldsymbol{r}-\boldsymbol{a}_s)]-2\psi(\boldsymbol{r})\Big\}$$

$$-2J\sum_{s=1}^{c}\delta_{\boldsymbol{r},\boldsymbol{a}_s}\psi(\boldsymbol{a}_s)$$

$$-2SJc_1\left\{\delta_{\boldsymbol{r},0}\sum_{s=1}^{c}\cos\frac{\boldsymbol{K}\cdot\boldsymbol{a}_s}{2}[\psi(\boldsymbol{a}_s)+\psi(-\boldsymbol{a}_s)]+2\sum_{s=1}^{c}\delta_{\boldsymbol{r},\boldsymbol{a}_s}\cos\frac{\boldsymbol{K}\cdot\boldsymbol{a}_s}{2}\psi(0)\right\} \tag{4·8b}$$

for $S\geq 1$.

The equation for the ideal spin-deviations is

$$E\psi(\boldsymbol{r})=-2SJ\sum_{s=1}^{c}\left\{\cos\frac{\boldsymbol{K}\cdot\boldsymbol{a}_s}{2}[\psi(\boldsymbol{r}+\boldsymbol{a}_s)+\psi(\boldsymbol{r}-\boldsymbol{a}_s)]-2\psi(\boldsymbol{r})\right\}, \tag{4·9}$$

of which solutions are

$$E_{\boldsymbol{K},k}=4SJ\sum_{s=1}^{c}\left(1-\cos\frac{\boldsymbol{K}\cdot\boldsymbol{a}_s}{2}\cos\boldsymbol{k}\cdot\boldsymbol{a}_s\right), \tag{4·10}$$

$$\psi_k(\boldsymbol{r})=L^{-1/2}e^{i\boldsymbol{k}\cdot\boldsymbol{r}}.$$

From these solutions we construct the Green function

$$G(\boldsymbol{r}-\boldsymbol{r}')=\sum_{k}\frac{\psi_k(\boldsymbol{r})\psi_k(\boldsymbol{r}')}{E-E_{\boldsymbol{K},k}} \tag{4·11}$$

satisfying

$$EG(\boldsymbol{r}-\boldsymbol{r}')=-2SJ\sum_{s=1}^{c}\left\{\cos\frac{\boldsymbol{K}\cdot\boldsymbol{a}_s}{2}[G(\boldsymbol{r}+\boldsymbol{a}_s-\boldsymbol{r}')+G(\boldsymbol{r}-\boldsymbol{a}_s-\boldsymbol{r}')]\right.$$

$$\left.-2G(\boldsymbol{r}-\boldsymbol{r}')\right\}-\delta(\boldsymbol{r}-\boldsymbol{r}').$$

By means of this function, the eigenvalue problems (4·8a, b) are transformed to

$$\psi(\boldsymbol{r})=\sum_{k}\frac{\psi_k(\boldsymbol{r})}{E-E_{\boldsymbol{K},k}}\left\{-2J\sum_{s=1}^{c}\psi_k(\boldsymbol{a}_s)\psi(\boldsymbol{a}_s)+v_0\psi_k(0)\psi(0)\right\} \tag{4·12a}$$

for $S=1/2$, and

$$\psi(\boldsymbol{r})=\sum_{k}\frac{\psi_k(\boldsymbol{r})}{E-E_{\boldsymbol{K},k}}\left\{-2J\sum_{s=1}^{c}\psi_k(\boldsymbol{a}_s)\psi(\boldsymbol{a}_s)\right.$$

$$\left.-4SJc_1\left[\psi_k(0)\sum_{s=1}^{c}\cos\frac{\boldsymbol{K}\cdot\boldsymbol{a}_s}{2}\psi(\boldsymbol{a}_s)+\sum_{s=1}^{c}\psi(\boldsymbol{a}_s)\cos\frac{\boldsymbol{K}\cdot\boldsymbol{a}_s}{2}\psi(0)\right]\right\} \tag{4·12b}$$

for $S\geq 1$. These equations are closed when we take up the equations for $\boldsymbol{r}=0$ and $\boldsymbol{a}_s$, $s=1, 2,\cdots, c$. If we write $\psi(0)$ in terms of $\psi(\boldsymbol{a}_s)$ as

$$v_0\psi(0)=2J\sum_{s=1}^{c}\cos\frac{\boldsymbol{K}\cdot\boldsymbol{a}_s}{2}\psi(\boldsymbol{a}_s) \tag{4·13a}$$

for $S=1/2$, and

$$(E-4SJc)\psi(0)=-4SJ(1+c_1)\sum_{s=1}^{c}\cos\frac{\boldsymbol{K}\cdot\boldsymbol{a}_s}{2}\psi(\boldsymbol{a}_s) \qquad (4\cdot 13\text{b})$$

for $S\geq 1$, which are obtained from (4·8a) and (4·8b) respectively, and eliminate $\psi(0)$ from c equations for $\boldsymbol{a}_s$, then we have

$$\psi(\boldsymbol{a}_s)=\sum_{k}\frac{\psi_k(\boldsymbol{a}_s)}{E-E_{\boldsymbol{K},k}}\left\{-2J\sum_{t=1}^{c}\psi_k(\boldsymbol{a}_t)\psi(\boldsymbol{a}_t)+2J\psi_k(0)\sum_{t=1}^{c}\cos\frac{\boldsymbol{K}\cdot\boldsymbol{a}_t}{2}\cdot\psi(\boldsymbol{a}_t)\right\} \qquad (4\cdot 14)$$

for $S\geq 1/2$, where the completeness of $\psi_k(\boldsymbol{r})$,

$$\sum_{k}\psi_k(\boldsymbol{a}_s)\psi_k(0)=0\,,$$

is used.

Our problem has thus been reduced to the one of solving the secular equation of a set of $c/2$ linear equations: for the solutions we are interested in are those for which $\psi(\boldsymbol{a}_s)=\psi(-\boldsymbol{a}_s)$. The coefficients $\sum_k\psi_k(\boldsymbol{r})\psi_k(\boldsymbol{r}')/(E-E_{\boldsymbol{K},k})$ may be able to be evaluated. However, the problem can be shown to be equivalent to that treated by Dyson, and so we do not here repeat the solution of our equations. The equivalence is easily seen by noticing that the equation (4·14) can be obtained from

$$\mathcal{H}_D=2SJ\sum_{<jk>}(a_j^*-a_k^*)(a_j-a_k)+J\sum_{<jk>}(a_j^*-a_k^*)^2a_j\,a_k$$

by the process just used to obtain Eq. (4·14) from $\mathcal{H}_G$ for the two-body case and that the complex conjugate of this $\mathcal{H}_D$ is the Hamiltonian adopted by Dyson; cf. eq. (57) of the first paper of Ref. 5. In this way we find that the operators for the ideal ferromagnet in Dyson's paper are nothing but the operators introduced by Holstein and Primakoff at least for two-body problem.

As the following processes, we can follow Dyson: we solve the scattering problem, calculate the energy shift for every set of $\boldsymbol{K}$ and $\boldsymbol{k}$ in first order perturbation, substitute it in Eqs. (4·1), (3·6) and (3·7), obtain the correction due to binary collisions to the free energy and the number of particles, then put $z=1$ and obtain the correction to N/L at $z=1$. We use this value in Eq. (2·10) and obtain the correction to the spontaneous magnetization to be of the order T^4 which is very small at low temperatures (see Ref. 5, p. 1243). As the result, we have obtained the results Dyson has given, by a method which is the generalization of that of Van Kranendonk, without meeting with the problems of non-diagonality, kinematical interactions or non-Hermitic Hamiltonians.

The binary collisions have been found to be unappreciable for very low temperatures. On the other hand, effects of collisions are expected to be important near the Curie point. To treat the latter region, we should apply a method for high density imperfect gases to our lattice gas. The author hopes that a powerful method for high density gases will appear and be applied to our lattice gas.

The author wishes to express his sincere thanks to Dr. K. Hiroike and Prof. H.

Ichimura for helpful discussions and criticisms.

References

1) T. Morita, Prog. Theor. Phys. **20** (1958), 614.
2) F. Bloch, Z. Physik **61** (1930), 206.
3) G. F. Koster and J. C. Slater, Phys. Rev. **95** (1954), 1167.
4) J. Van Kranendonk, Physica **21** (1955), 749 and 925.
5) F. J. Dyson, Phys. Rev. **102** (1956), 1217 and 1230.
6) T. Holstein and H. Primakoff, Phys. Rev. **58** (1940), 1098.
7) T. Morita, Prog. Theor. Phys. **18** (1957), 462.
8) H. D. Ursell, Proc. Camb. Phil. Soc. **23** (1927), 685.
e. g. J. O. Hirschfelder, C. F. Curtiss and R. Byron Bird, *Molecular Theory of Gases and Liquids* (Wiley, 1954), 138.

Progress of Theoretical Physics, Vol. 20, No. 5, November 1958

Notes on the Magnetic Mirror Effect*, **

Masatomo SATO

Department of Physics, College of Science and Engineering, Nihon University, Tokyo

(Received August 7, 1958)

An elementary theory is presented to estimate the rate of charged particle loss from the ends of a magnetic mirror device.

Introduction

One of the simplest ways of the magnetic confinement of hot plasmas is to construct a parallel magnetic wall around the plasma column. Although this type of configuration has the advantage of having no dynamical instabilities such as those present in the pinch effect, it has yet such an essential disadvantage that the charged particles can escape freely from the confining region when the magetic field lines are straight and the region has finite ends. One method of making up for this disadvantage is to use an endless apparatus, of doughnut type, say, and another is to make the field lines squeezed at the both ends to make use of the magnetic mirror effect[1] to prevent the particle loss.

In the former case, one has to make some special devices of applying other magnetic fields, since one can get no static equilibrium configuration only with the field lines parallel to the plasma column axis.[2] In this note we shall consider the latter case to see how the magnetic mirror effect can reduce the rate of particle loss.

§ 1. The mirror condition

Let the magnetic field be static, and assume that the motion of the charged particles in the plasma can be described by the first order orbit theory.[1),3)] It is well known that this theory can be applied with a good accuracy when the spatial change of the field over a distance of the order of the particle gyro-radius is sufficiently smaller than the field strength itself.*** It will further be necessary that the go-and-return time of a particle between the mirrors is sufficiently smaller than the interparticle collision time,

---

* Lectured at the Plasma Physics Symposium held at the Institute for Fundamental Physics, Kyoto University, during 7th-15th May, 1958.

** After this paper had been submitted to the editor, the recent issue of the Nuovo Cimento came to hand, in which E. Persico and J. G. Linhart[5] are treating the "Plasma Loss from Magnetic Bottles" by a similar way as used in this paper.

*** In other words, when the gyro-radius is sufficiently smaller than the linear dimension of the confinement region.

since otherwise the system would have to be treated something like a fluid, rather than by the one particle picture. However, this condition may be well satisfied in a high temperature (and low density) plasma of interest, since the go-and-return time is proportional to the inverse square root, and the collision time to the three-half powers, of the plasma temperature (cf. § 3·1).

Under this assumption, the particle undergoes a spiral motion along the field line. Let the pitch angle of the spiral be θ at the middle part of the confining region, and the field strength be B_0 and B_m at the middle part and ends, respectively. Then the particle spiraling towards the end will be reflected back if θ is larger than the critical angle θ_c given by[1]

$$\sin^2\theta_c = B_0/B_m. \tag{1·1}$$

We can then divide the particles into two groups according to the magnitude of θ: the reflecting group for which $\theta > \theta_c$ and the escaping group for which $\theta < \theta_c$. This corresponds to the division of the velocity space, associated with each point of the middle

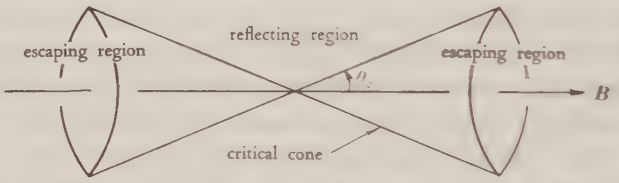


Fig. 1. Escaping and reflecting regions, considered in the velocity space associated with each point of the middle part of the confining region.

part, into the reflecting and the escaping regions, as shown in Fig. 1. Here the angular spread of the critical cone is given by the solid angle

$$\Omega_c = 2\pi(1-\cos\theta_c) \approx \pi(B_0/B_m) \tag{1·2}$$

with $(B_0/B_m) \ll 1$.

Actually, those particles which initially belonged to the reflecting group may also be scattered into the escaping group by the effect of inter-particle collisions, and the confinement by mirrors lasts only a finite time. To estimate the rate of the particle loss is the purpose of the following section.

§ 2. The rate of particle loss

The particle density (of positive ions, say) may be divided into two parts according to the division mentioned in § 1:

$$n = n_r + n_e, \tag{2·1}$$

where the meaning of the suffices will be obvious. The rate of change in time of n_r and n_e are represented by the mixing rates of both parts caused by scattering and by the net decrease of n due to the particle loss through the mirrors:

$$dn_r/dt = -w(r\to e)\,n_r + w(e\to r)\,n_e \qquad (2\cdot 2)$$

$$dn_e/dt = +w(r\to e)\,n_r - w(e\to r)\,n_e - (\partial n_e/\partial t)_{esc}. \qquad (2\cdot 3)$$

Here $w(r\to e)$ means the probability per unit time for a particle of the reflecting group to go into the escaping group as a result of collisions, and $w(e\to r)$ is a similar probability for the inverse process. The total density n decreases according to

$$dn/dt = -(\partial n_e/\partial t)_{esc}. \qquad (2\cdot 4)$$

For simplicity, we shall assume here that those particles which suffer the large angle deflections go into respective groups with the relative probabilities proportional to the solid angles for the respective groups. This assumption will be natural if the encounters of particles occur with sufficient randomness; but further consideration may be necessary on this point.

Taking the above assumption for granted, eqs. (2·2) and (2·3) are rewritten as follows. Let us define ω by

$$\omega = 2\Omega_c/4\pi \approx \tfrac{1}{2}(B_0/B_m) \qquad (2\cdot 5)$$

and let γ_s be the collision frequency for the large angle reflections (including, of course, those large angle ones as a cumulative result of many small angle scatters). Then

$$dn_r/dt = -\gamma_s\omega\, n_r + \gamma_s(1-\omega)\,n_e, \qquad (2\cdot 6)$$

$$dn_e/dt = +\gamma_s\omega\, n_r - \gamma_s(1-\omega)\,n_e - (\partial n_e/\partial t)_{esc}. \qquad (2\cdot 7)$$

In solving these differential equations, let us assume that γ_s is nearly constant. Further we assume that

$$(\partial n_e/\partial t)_{esc.} = \gamma_e n_e, \qquad (2\cdot 8)$$

where γ_e is the reciprocal of the time for a particle of the escaping group to leave the confining region. Then, eqs. (2·6) and (2·7) become:

$$d^2n_i/dt^2 + (\gamma_s+\gamma_e)(dn_i/dt) + \omega\gamma_s\gamma_e n_i = 0. \quad (i=r,\ e) \qquad (2\cdot 9)$$

The suitable initial conditions will be

$$\begin{cases} n_r(t=0) = n_r^{(0)}, & (dn_r/dt)_{t=0} = 0\,; \qquad (2\cdot 10) \\ n_e(t=0) = n_e^{(0)}, & (dn_e/dt)_{t=0} = -\gamma_e n_e^{(0)}, \qquad (2\cdot 11) \end{cases}$$

meaning also that the initial total density $n^{(0)}$ is divided into both groups proportional to their solid angles: $n_e^{(0)} = \omega n^{(0)}$, $n_r^{(0)} = (1-\omega)\,n^{(0)}$. $(2\cdot 12)$

Then the solutions of eq. (2·8) are:

$$n_r = \frac{n_r^{(0)}}{\lambda^{(+)}-\lambda^{(-)}}\left\{\lambda^{(+)}e^{-\lambda^{(-)}t} - \lambda^{(-)}e^{-\lambda^{(+)}t}\right\} \qquad (2\cdot 13)$$

$$n_e = \frac{n_e^{(0)}}{\lambda^{(+)}-\lambda^{(-)}}\left\{(\lambda^{(+)}-\gamma_e)\,e^{-\lambda^{(-)}t} - (\lambda^{(-)}-\gamma_e)\,e^{-\lambda^{(+)}t}\right\} \qquad (2\cdot 14)$$

$$n=n_r+n_e=\frac{n^{(0)}}{\lambda^{(+)}-\lambda^{(-)}}\{(\lambda^{(+)}-\omega\gamma_e)\,e^{-\lambda^{(-)}t}-(\lambda^{(-)}-\omega\gamma_e)\,e^{-\lambda^{(+)}t}\},\qquad(2\cdot15)$$

where

$$\lambda^{(\pm)}=\tfrac{1}{2}\{(\gamma_s+\gamma_e)\pm\sqrt{(\gamma_s+\gamma_e)^2-4\omega\gamma_s\gamma_e}\}.\qquad(2\cdot16)$$

It can be seen easily that $\lambda^{(-)}$ is always much larger than $\lambda^{(+)}$ and the term proportional to $\exp(-\lambda^{(+)}t)$ can be neglected for $t\gg(\lambda^{(+)})^{-1}$.

We are interested in the case of $\gamma_e\gg\gamma_s$. Using $\omega\ll1$, we get the following asymptotic solution for $t\gg(\lambda^{(+)})^{-1}$:

For $\gamma_e\gg\gamma_s$*: $\quad n\approx n_r\approx n^{(0)}e^{-\omega\gamma_s t},\quad n_e\approx\omega n^{(0)}(\gamma_s/\gamma_e)\,e^{-\omega\gamma_s t}$; $\qquad(2\cdot17)$

For $\gamma_e=\gamma_s$: $\quad n\approx n_r\approx n^{(0)}e^{-\omega\gamma_s t/2},\quad n_e\approx\omega(n^{(0)}/2)\,e^{-\omega\gamma_s t/2}$. $\qquad(2\cdot18)$

Thus the time scale for the particle loss, t_L, is found to be

$$t_L\approx t_s/\omega\approx 2t_s(B_m/B_0),\qquad(2\cdot19)$$

where $t_s=(\gamma_s)^{-1}$ is the collision time, which may be identified with the mean collision time for positive ions, since the main rate of particle loss will be determined by the heavier particles (for numerical values of t_s, cf. § 3·1). That the loss rate is determined by the longer one of the two characteristic times t_s and $t_e(=\gamma_e^{-1})$, may be understood qualitatively. For in the case of $t_e\ll t_s$, for instance, the particles having once belonged to the escaping group will leave the confining region at once, while it takes a longer time for a particle to move from the reflecting to the escaping group. (Therefore, only very few particles are present in the escaping group in this case.) For $t_e\gg t_s$, on the other hand, the mixing of two groups due to the collisions becomes important, bringing about the division of particles into two groups always in proportion to the solid angles. In this case it is legitimate to think that the plasma diffuses as a whole through the mirror holes, the aperture of which is given by ω (see equations in the footnote to eq. (2·17), above). It should be noted also that the result $t_L\propto\omega^{-1}$, obtained in eq. (2·19), is essentially dependent on the assumption which has been made earlier below eq. (2·4).

§ 3. Remarks

3·1. The escaping time t_e, that is, the time for a particle of the escaping group to leave the confining region, is given by

$$t_e=\int_0^l dx/v_{\parallel}(x),\qquad(3\cdot1)$$

where $v_{\parallel}(x)$ is the velocity component parallel to the field line, along which the length x from the middle point is measured, and l is the total length of the confining region. Now, $v_{\parallel}(x)$ is easily computed from the conservation law of the kinetic energy of the particle and the adiabatic invariance theorem of the gyro-magnetic moment, i. e.,

* For $\gamma_e\ll\gamma_s$, we get $n_i\sim n_i^{(0)}e^{-\omega\gamma_s t}$ for $i=e$ and r.

$$v_{\parallel}(x)=v_0\{1-(v_{0\perp}/v_0)^2 B(x)/B_0)\}^{1/2}, \qquad (3\cdot 2)$$

where v_0 is the speed at the middle part. Similarly, the reflecting time (or, the go-and-return time, as we called it before) t_r for a particle of the reflecting group is

$$t_r=2\int_0^{l_r} dx/v_{\parallel}(x), \qquad (3\cdot 3)$$

where $x=l_r$ is the position at which the right-hand side of eq. (3·2) vanishes.

It can be shown from a simple consideration, however, that the magnitude of t_e and t_r given by eqs. (3·1) and (3·3) are both of the order of $l/v_{0\parallel}$, which is the escaping time when the magnetic field is uniform (for t_r, we assumed that $l_r\sim l$). That is, the middle region where $v_{\parallel}(x)\approx v_{0\parallel}$ ($\equiv v_{\parallel}(0)$) gives a main contribution to t_e and t_r. Therefore,

$$t_e\approx t_r\approx l/(kT/m)^{1/2}, \qquad (3\cdot 4)$$

where k is the Boltzmann constant, T the temperature, and m is the particle mass.

Let us compare t_e and t_r with the collision time t_s, where the latter may be represented by the so-called self-collision time, t_c. This is given by (cf. eq. (5–26) of reference[1]) :

$$t_c=m^{1/2}(3kT)^{3/2}/(8\times 0.714\pi ne^4 Z^4\ln\Lambda) \qquad (3\cdot 5)$$

$$=16.1\,(T^{3/2}/n\ln\Lambda) \quad \text{(for D ions)}. \qquad (3\cdot 5')$$

Then from eqs. (3·4) and (3·5),

$$t_e/t_s\approx t_r/t_s\approx 10^{-5}\ln\Lambda(nl/T^2)\approx 10^{-4}(nl/T^2). \qquad (3\cdot 6)$$

(T : in °K, l : in cm, n : in particles·cm^{-3})

Thus, for conditions usually considered in the controlled thermonuclear fusion research ($l\sim 10^2$ cm, $n\sim 10^{12-15}$, $T\sim 10^{6-9}$), we can choose the set of parameters to let $t_{e,r}/t_s\ll 1$, as mentioned in § 1. Using eq. (2·19), we may see then that the particles between the mirrors repeat a number of go-and-return motions before they are lost. Here, the rather strong temperature dependence of t_l is to be noted : Higher temperature makes the confinement by mirrors easier.

3·2.—The particle loss considered above is the one which is mainly caused by the collisions between positive ions. The ion-electron collisions, on the other hand, mainly determine the rates of plasma loss across the field lines and the field penetration into the plasma. These effects are macroscopically represented by the finite electrical resistivity transverse to the magnetic field, and are estimated by Spitzer (cf. eqs. (3–16) and (5–39) of reference 1)). Rough estimate of the transverse diffusion time (t_D) and the field penetration time (t_B) may be expressed as

$$t_D\approx \bar{B}^2 R^2/p\eta, \quad t_B\approx 4\pi R_0^2/\eta, \qquad (3\cdot 7)$$

where $\bar{B}$ is the mean value of the field, p the plasma pressure, η the electrical resistivity

in e. m. u., and R and R_0 are the radii of the confinement region and the plasma region, respectively.

We shall first compare t_B with t_L, eq. (2·19). For practical purposes, it will be reasonable to require $t_L \gtrsim t_B$, since, then, the ohmic heating of plasma will becomes effective. Using $\eta = 1.39 \times 10^{13} \ln \Lambda \, T^{-3/2}$ e. m. u. as given by eq. (5–39) of 1) we get, as the equivalent condition to $t_L \gtrsim t_B$,

$$B_m / B_0 \gtrsim 6 R_0^2 (n/10^{14}). \tag{3·8}$$

For a low density plasma with the confinement radius of several cm, the realization of the condition (3·8) will not be technically impossible.

On the other hand, t_D will be much larger than t_B, because R is evidently larger than R_0 and p is smaller than the magnetic pressure $\bar{B}^2/8\pi$ in the quasi-equilibrium state where the fields are partly penetrated into the plasma. Therefore, t_D may be considered as larger than t_L, and the diffusion loss at right angles to the field lines may be neglected in comparison with the loss through mirror ends.

3·3.—A possible trouble inherent in the simple mirror bottle is the dynamical (flute-type) instabilities of the plasma, which may be present when the plasma boundary is convex along the field lines,[4] although this possibility may be reduced when we use a sufficiently long bottle and if the field is partially penetrated in the plasma. If such an instability proved to be serious,* an obvious way of remedying this situation will be to let the field lines convex everywhere against the plasma by using a quadrupole magnetic field. In such a case, the rate of particle loss may also be estimated in a similar way as treated above in § 2.[5]

This work was initially motivated through the discussions with Prof. C. Hayashi (Kyoto University), Prof. G. Tominaga and Mr. H. Onishi. The author wishes to express his thanks to them and to the members of the Nuclear Fusion Research Group of Nihon University for helpful discussions.

Note added in proof Instead of putting $w(e \to r) = \gamma_s (1 - \omega)$ as used in eq. (2·6), one would have to put $w(e \to r) = \gamma'$ as adopted in 5), where γ' is the reciprocal of the the time between two collisions at an angle larger than θ_c and approximately given by[5] (several) $\cdot (n/\omega T^{3/2})$. No essential change is necessary for the subsequent argument but the initial conditions (2·10) and (2·11) should be slightly modified for $(dn, dt)_{t=0}$ $(i = r, e)$ and the time scale for the particle loss (t_I) becomes $1 + (\gamma'/\gamma_e)$ times as large as the one given by eq. (2·19), with $\gamma'/\gamma_e \sim$ (several) $\cdot (10^{-4} n l/\omega T^2)$.

References

1) L. Spitzer, *Physics of Fully Ionized Gases*, Interscience Publishers, New York, 1956.
2) See, e. g., L. Biermann und A. Schlüter, Z. Naturforschg. **12a** (1957), 805.
3) H. Alfvén, *Cosmical Electrodynamics*, Oxford, 1950.
4) M. N. Rosenbluth and C. L. Longmire, Annals of Physics, **1** (1957), 120.
5) E. Persico and J. G. Lnihart, Nuovo Cimento 8 (1958), 740.

* After the manuscript of this paper had been prepared, we had the chance of reading Dr. R. F. Post's report about Livermore's Mirror Machine, at the 1958 Spring Meeting of the American Physical Society. In this report, he mentioned that the plasma in the mirror configuration was found to be stable experimentally

Progress of Theoretical Physics, Vol. 20, No. 5, November 1958

A General Theory of Magnetic Double Resonance*

Kazuhisa TOMITA

Department of Physics, Faculty of Science, University of Kyoto, Kyoto

(Received July 21, 1958)

By using a similar method as was used in the preceding paper, it is proposed to present a general theory for describing a system which consists of two interacting different species of spin, one of them being saturated by a strong resonant radiation field and the other being detected by a weak resonant radiation field. Two rather different extreme cases are classified according to the relative magnitude of the static strength of the mutual interaction σ_0 and the characteristic frequency of fluctuation ϕ_0 of the relevant environment. A) When the fluctuation of the environment dominates ($\phi_0 \gg \sigma_0$), then we are led to the situation which may be called a generalized "Overhauser effect", i. e. the process of saturation works as an energy pumping. Without the effect of the continuous saturation the simultaneous equation reduces to a generalized Solomon equation originally proposed for describing the coupled free induction of two kinds of nuclear spin system. B) If, on the other hand, the static strength of the mutual interaction dominates over the fluctuation ($\sigma_0 \gg \phi_0$), then there results a phenomenon called "Saturational narrowing", i. e. the process of saturation works as a local field modulation. This latter case has also been illustrated by an application to the double resonance experiment on thallous fluoride crystal.

§ 1. Introduction

We suppose that there are two species of spin system having different magnetic moments, interacting with each other, and being immersed in a strong constant magnetic field. When we saturate one kind of spin by a strong resonant radiation, and at the same time trace the behaviour of the other kind of spin by using a weak resonant radiation, the method is called "Double resonance". As a result of this type of experiment, two markedly different situations are found. The first kind of situation is that of the so-called "Overhauser polarization effect",[1)~9)] and its typical example is found in alkali metals. The effect has been observed[3)] as a remarkable enhancement of the nuclear magnetic resonance intensity under the application of a strong radiofrequency field which is resonant with the precession of the electron spins. A phenomenon of the same type has been found by Solomon[10)] for two different nuclear spin species, i. e. for liquid hydrogen fluoride, although the effect is much smaller quantitatively. However, a completely different situation is found for the nuclear magnetic double resonance in ionic crystals. When the resonance is saturated for one species of nucleus in these crystals, the width of the resonance absorption of the other species is considerably decreased, which should be called "Saturational narrowing". Saito[11)] has observed this type of phenomenon in TlF and other

* Read at the Hiroshima Meeting of the Physical Society of Japan on April 6, 1957.

substances, and Searls and Cotts[12] observed the effect in NaF crystal. The corresponding effect has been found in the non-steady, echo-type experiment of Herzog and Hahn[13] on $NaClO_3$ crystal. The main purpose of the present article is to clarify under what conditions the two markedly different situations result. It will be shown: 1) When the fluctuation of the environment — the system outside the two species of spin system, which is conventionally called the lattice system — is strong enough to supersede the coupling between two species of spin, a general Overhauser polarization effect is expected. 2) In contrast, when the intensity of coupling dominates over the lattice fluctuation, there results the saturational narrowing. This latter situation is illustrated by the case of TlF.

§ 2. Fundamental equation

In a previous paper[14]* the present writer discussed the behaviour of a system of single spin species under a strong resonant radiation by starting with the differential equation which is obeyed by the density matrix of the whole system. Here we use essentially the same method and adopt similar notations and definitions for the purpose of convenience. When there are two kinds of spin system, which henceforth will be called the s-system and the t-system, there can be several different kinds of relaxation mechanism, that is the relaxation which should survive even when the interaction G between the two systems is virtually removed, and the relaxation which comes in through the existence of the interaction G itself. As the first kind of process was treated in the previous paper in detail, we here concentrate our attention on the second kind. This process is determined essentially by the interaction G (Hamiltonian divided by $\hbar$), which is given by

$$G=\sum_s\sum_t G_{st}=\sum_s\sum_t\sum_{\mu_s}\sum_{\mu_t} I_s^{\mu_s} I_t^{\mu_t} F_{st}^{\bar{\mu}_s \bar{\mu}_t} \quad (|\mu_s+\mu_t|=0,\ 1,\ 2). \tag{2·1}$$

Here I_s and I_t denote the spin operators of the s- and the t-system respectively, and μ_s, μ_t stand for the number of respective quantum jumps (The bar on the top stands for the minus sign). The density matrix describing the whole system obeys the differential equation

$$\frac{d\rho}{dt}=\frac{i}{\hbar}[\rho,\ \mathcal{H}], \tag{2·2}$$

where the total Hamiltonian $\mathcal{H}$ is given by

$$\mathcal{H}=\hbar(H_s+H_t+G+F). \tag{2·3}$$

Here we used the abbreviations

$$H_s=E_s+D_s+G_s+F_s, \tag{2·4}$$

$$H_t=E_t+D_t+G_t+F_t, \tag{2·5}$$

where E stands for the Zeeman energy associated with the constant field H_0, i.e.

* Henceforth this paper will be referred to as paper I.

$$\begin{cases} E_s = -\gamma_s H_0 \sum_s I_{sz} = -\omega_s \sum_s I_{sz}, & (\omega_s = \gamma_s H_0) , \\ E_t = -\gamma_t H_0 \sum_t I_{tz} = -\omega_t \sum_t I_{tz}, & (\omega_t = \gamma_t H_0) . \end{cases} \tag{2·6}$$

$$\tag{2·7}$$

D stands for the energy associated with the interaction with the strong resonant component $H_1 e^{\mp i\omega t}$ of the radio-frequency field,

$$\begin{cases} D_s = -\omega_{1s} \sum_s (I_{sx} \cos\omega t - I_{sy} \sin\omega t), & (\omega_{1s} = \gamma_s H_1), \\ D_t = -\omega_{1t} \sum_t (I_{tx} \cos\omega t - I_{ty} \sin\omega t), & (\omega_{1t} = \gamma_t H_1). \end{cases} \tag{2·8}$$

$$\tag{2·9}$$

G_s and G_t are responsible for the first kind relaxation and F_s and F_t are the relevant lattice degrees of freedom for the s- and the t- system respectively. G is the interaction between the two systems and F is the relevant lattice degree of freedom.

Note that, with superfixes, $F_{st}^{\mu_s \mu_t}$ stands for the part of G, which is dependent on the lattice F. (cf. (2·1))

In order to take the effect of saturation fully into account, first we change the frame of reference to the one which is rotating with the rotating field H_1 by way of the transformation

$$A^T = e^{iE't} A e^{-iE't}, \tag{2·10}$$

where

$$E' = -\omega (\sum_s I_{sz} + \sum_t I_{tz}). \tag{2·11}$$

Then the equation of motion in the rotating frame becomes

$$\frac{d\rho^T}{dt} = i[\rho^T,\ E_s^{Te} + G_s^T + F_s + E_t^{Te} + G_t^T + F_t + G^T + F], \tag{2·12}$$

where $E_s^{Te}(E_t^{Te})$ denotes the Zeeman energy of the s-system (t-system), i.e.

$$\begin{cases} E_s^{Te} = -\{\Delta_s \sum_s I_{sz} + \omega_{1s} \sum_s I_{sx}\}\ ; & (\Delta_s \equiv \omega_s - \omega), \\ E_t^{Te} = -\{\Delta_t \sum_t I_{tz} + \omega_{1t} \sum_t I_{tx}\}\ ; & (\Delta_t \equiv \omega_t - \omega). \end{cases} \tag{2·13}$$

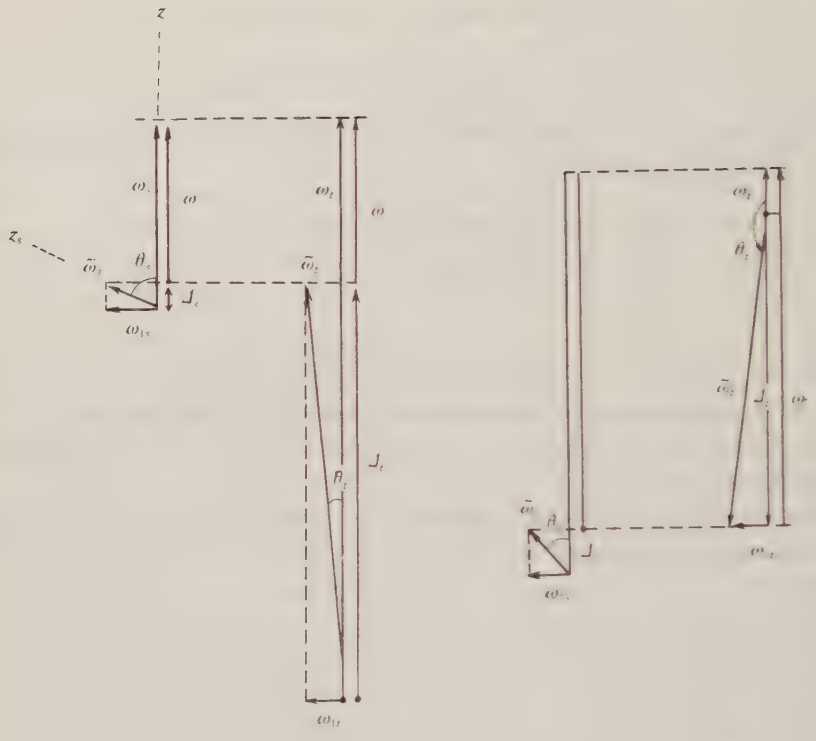
$$\tag{2·14}$$

When we choose $\omega \simeq \omega_s$ the effective field $\tilde{\omega}_s$ for the s-system in the rotating frame becomes nearly perpendicular to ω_s, whereas the angle between ω_t and the effective field becomes either $\simeq 0$ or $\simeq \pi$. As we have a condition $\Delta_t \gg \omega_{1t}$, we may safely neglect ω_{1t} in the following. The G's can be decomposed into partial matrices in the rotating frame as follows.

$$G_s^T = \sum_{\mu_s} I^{\mu_s} F_s^{\bar{\mu}_s} e^{i\omega_{\mu_s} t}, \quad (\omega_{\mu_s} = -\mu_s \omega) \tag{2·15}$$

$$G_t^T = \sum_{\mu_t} I^{\mu_t} F_t^{\bar{\mu}_t} e^{i\omega_{\mu_t} t}, \quad (\omega_{\mu_t} = -\mu_t \omega) \tag{2·16}$$

$$G_{st}^T = \sum_{\mu_s} \sum_{\mu_t} I^{\mu_s} I^{\mu_t} F_{st}^{\bar{\mu}_s \bar{\mu}_t} e^{i(\omega_{\mu s} + \omega_{\mu t}) t}. \tag{2·17}$$



(a) (b)

Fig. 1. The description in the rotating system of coordinates.
(a) The case in which $\omega_s \ll \omega_t$.
(b) The case in which $\omega_s \gg \omega_t$.

Here and in what follows we dispense with the suffixes s and t where there is no danger of confusion.

Under the assumption that our system is essentially paramagnetic, the part described by G's is considered to be a small perturbation as compared with the remaining part of the Hamiltonian. Therefore, we further transform to the interaction representation by way of the operation

$$A^* = e^{i(E_s{}^T{}^e + F_s + E_t{}^T{}^e + F_t + F)t} A e^{-i(E_s{}^T{}^e + F_s + E_t{}^T{}^e + F_t + F)t}. \tag{2·18}$$

The density ρ^{T*} in the new frame obeys the equation

$$\frac{d\rho^{T*}}{dt} = i[\rho^{T*},\ G_s^{T*}(t) + G_t^{T*}(t) + G^{T*}(t)], \tag{2·19}$$

where

$$\begin{aligned} G_s^{T*} &= \sum_{\lambda_s}\sum_{\mu_s} I^{\mu_s *} F_s^{\bar{\mu}_s}(\lambda_s)\, e^{i(\omega_{\lambda_s} + \omega_{\mu_s})t} \\ &= \sum_{\lambda_s}\sum_{\mu_s}\sum_{\nu_s} a_{\mu_s \nu_s} I(\nu_s) F_s^{\bar{\mu}_s}(\lambda_s)\, e^{i(\omega_{\lambda_s} + \omega_{\mu_s} + \omega_{\nu_s})t}, \end{aligned} \tag{2·20}$$

$$G_t^{T*}=\sum_{\lambda_t}\sum_{\mu_t}\sum_{\nu_t} a_{\mu_t\nu_t} I(\nu_t) F_t^{\bar{\mu}_t}(\lambda_t)\, e^{i(\omega_{\lambda_t}+\omega_{\mu_t}+\omega_{\nu_t})t}, \tag{2·21}$$

$$G_{st}^{T*}=\sum_{\lambda}\sum_{\mu_s}\sum_{\nu_s}\sum_{\mu_t}\sum_{\nu_t} a_{\mu_s\nu_s} a_{\mu_t\nu_t} I(\nu_s) I(\nu_t) F_{st}^{\bar{\mu}_s\bar{\mu}_t}(\lambda)\, e^{i(\omega_\lambda+\omega_{\mu_s}+\omega_{\nu_s}+\omega_{\mu_t}+\omega_{\nu_t})t}. \tag{2·22}$$

In the above equations we have used the partial matrix decompositions like

$$F^{\lambda *} \equiv e^{iFt} F^{\mu} e^{-iFt} = \sum_{\lambda} F^{\mu}(\lambda)\, e^{i\omega_\lambda t}, \tag{2·23}$$

$$I(\nu_s)^* \equiv e^{iE_s T^e t} I(\nu_s)\, e^{-iE_s T^e t} = I(\nu_s)\, e^{i\omega_{\nu_s} t}, \tag{2·24}*$$

$$I(\nu_t)^* \equiv e^{iE_t T^e t} I(\nu_t)\, e^{-iE_t T^e t} = I(\nu_t)\, e^{i\omega_{\nu_t} t}, \tag{2·25}*$$

where a new z'-axis has been chosen along the direction of the effective field for the s-system (or t-system) in the rotating frame, so that we have obtained the simple form of decomposition (2·24) (or (2·25)). This additional change of frame is effected by a transformation like

$$\tilde{I}^{\mu} \equiv e^{i\theta I_y} I^{\mu} e^{-i\theta I_y} = \sum_{\nu} a_{\mu\nu} I(\nu), \tag{2·26}$$

or explicitly by

$$\begin{pmatrix} \tilde{I}^0 \\ \tilde{I}^+ \\ \tilde{I}^- \end{pmatrix} = \begin{pmatrix} \cos\theta & -\dfrac{\sin\theta}{2} & -\dfrac{\sin\theta}{2} \\ \sin\theta & \dfrac{\cos\theta+1}{2} & \dfrac{\cos\theta-1}{2} \\ \sin\theta & \dfrac{\cos\theta-1}{2} & \dfrac{\cos\theta+1}{2} \end{pmatrix} \begin{pmatrix} I(0) \\ I(+) \\ I(-) \end{pmatrix}, \tag{2·27}$$

where θ should be chosen as θ_s or θ_t for the s- and the t-system respectively. We have also used the abbreviations

$$\omega_{\nu_s} = -\nu_s \tilde{\omega}_s = -\nu_s \sqrt{\omega_{1s}^2 + \Delta_s^2}, \tag{2·28}$$

$$\omega_{\nu_t} = -\nu_t \tilde{\omega}_t = -\nu_t \sqrt{\omega_{1t}^2 + \Delta_t^2}. \tag{2·29}$$

If we write down the short time development of the density according to the equation (2·19), we have

$$\begin{aligned}\rho^{T*}(t) &= \rho^{T*}(0) + \Delta_1 \rho^{T*}(0) + \Delta_2 \rho^{T*}(0) + \cdots\cdots \\ &= \rho^{T*}(0) - i\int_0^t dt_1 \sum_s [G_s^{T*}(t_1),\ \rho^{T*}(0)] - \int_0^t dt_1 \int_0^{t_1} dt_2 \sum_{s,s'} [G_s^{T*}(t_1)[G_{s'}^{T*}(t_2),\ \rho^{T*}(0)]] \\ &\quad - i\int_0^t dt_1 \sum_t [G_t^{T*}(t_1),\ \rho^{T*}(0)] - \int_0^t dt_1 \int_0^{t_1} dt_2 \sum_{t,t'} [G_t^{T*}(t_1)[G_{t'}^{T*}(t_2),\ \rho^{T*}(0)]]\end{aligned}$$

* In the corresponding equation (1·13) of paper I $\sum_{\nu}$ should be erazed in the last expression.

$$-i\int_0^t dt_1[G^{T*}(t_1),\ \rho^{T*}(0)]-\int_0^t dt_1\int_0^{t_1} dt_2[G^{T*}(t_1)[G^{T*}(t_2),\ \rho^{T*}(0)]]+\cdots\cdots \tag{2·30}$$

under the assumption that the three different lattice degrees of freedom, i.e. F_s, F_t, and F_c do not interfere with each other. As the part of relaxation corresponding to the isolated s- or t-system has already been treated in the previous paper, we here treat only the part containing the interaction G in detail. Let us suppose that the lattice degree of freedom F is so large that the density matrix characterizing its state is practically independent of the spin complexion, then the total density matrix ρ can be factorized into those of the spin system σ and the lattice system P_F

$$\rho=\sigma\times P_F \tag{2·31}$$

where P_F may be assumed as the canonical distribution with respect to the free lattice F, i. e.

$$P_F=\exp\beta(\Omega-\hbar F)\,; \tag{2·32}$$

$$\beta=1/kT \quad \text{and} \quad \exp-\beta\Omega\equiv\mathrm{Tr}\,[\exp-\beta\hbar F].$$

If we take an average of the initial condition over the lattice states, then we have the density matrix describing the spin system as follows,

$$\sigma^{T*}(t)=\sigma^{T*}(0)-i\int_0^t dt_1[\langle G^{T*}(t_1)P_F\rangle_F,\ \sigma^{T*}(0)]$$

$$-\int_0^t dt_1\int_0^{t_1} dt_2\langle[G^{T*}(t_1)[G^{T*}(t_2),\ \sigma^{T*}(0)P_F]]\rangle_F+\cdots\cdots, \tag{2·33}$$

where $\langle\ \ \rangle_F$ indicates the operation of taking partial trace with respect to the lattice system F.

At this stage we further introduce a kind of Hartree approximation and decompose the spin density into those of the s-system and the t-system. This is permissible so far as there exists a condition

$$\hbar|\omega_s-\omega_t|\gg|G|\,, \tag{2·34}$$

because in this case only the time average of the effect of interaction can be observed. Essentially the same situation will exist so long as $|\omega_s-\omega_t|>\omega_{1/2}$, therefore we put

$$\sigma^{T*}=\sigma_s^{T*}\times\sigma_t^{T*}. \tag{2·35}$$

In this approximation it is possible to talk about the equation of motion which is obeyed by a single kind of spin system. For simplicity we assume that spins of the same kind are equivalent with respect to the geometrical configuration of their environment. Let us derive, for example, the equation of motion obeyed by the t-system by taking an average over the fluctuation of the s-system. Corresponding to the situation (a) in Fig. 1, we have $\theta_t\simeq 0$, $a_{\mu_t\nu_t}\simeq\delta_{\mu_t\nu_t}$, $\tilde{\omega}_t\simeq\Delta_t=\omega_t-\omega$, and

$$G_{st}^{T*}(t)=\sum_{\lambda}\sum_{\mu_s}\sum_{\mu_t}I^{\mu_s*}I^{\mu_t}F_{st}^{\bar{\mu}_s\bar{\mu}_t}(\lambda)\,e^{i(\omega_\lambda+\omega_{\mu_s}+\omega^0_{\mu_t})t}, \tag{2·36}$$

where

$$\omega^0{}_{\mu_t}=-\mu_t\omega_t.$$

Corresponding to the situation (b) in Fig. 1, we have $\theta_t\simeq\pi$, $a_{\mu_t\nu_t}=-\delta_{\mu_t\nu_t}$, $\tilde{\omega}_t\simeq\omega-\omega_t$, and

$$G_{st}^{T*}(t)=-\sum_{\lambda}\sum_{\mu_s}\sum_{\mu_t}I^{\mu_s*}I^{\mu_t}F_{st}^{\bar{\mu}_s\bar{\mu}_t}(\lambda)\,e^{i(\omega_\lambda+\omega_{\mu_s}+\omega^0_{\mu_t})t}. \tag{2·37}$$

Let us confine ourselves to the case (a), $\theta_t\simeq 0$, in the following. Then, after taking an average over the state of s-system, we have

$$\sigma_t^{T*}(t)=\sigma_t^{T*}(0)+\Delta_1\sigma_t^{T*}+\Delta_2\sigma_t^{T*}+\cdots\cdots, \tag{2·38}$$

of which the first order term is given by

$$\Delta_1\sigma_t^{T*}=-i\int_0^t dt_1[\sum_s\langle G_{st}^{T*}(t_1)\rangle_{F,s},\ \sigma_t^{T*}(0)], \tag{2·39}$$

where

$$\langle G_{st}^{T*}(t_1)\rangle_{F,s}=\sum_{\mu_s}\sum_{\mu_t}e^{i(\omega_{\mu_s}+\omega_{\mu_t})t_1}\langle F_{st}^{\bar{\mu}_s\bar{\mu}_t}(0)\rangle_F\langle I^{\mu_s*}\rangle_S I^{\mu_t} \tag{2·40}$$

from (2·36), and we have used the convention $\langle A\rangle_s=tr(\sigma_s^{T*}A)$. By definition $\langle I^{\mu_s*}\rangle$ vanishes except for $\mu_s=0$, and only $\mu_t=0$ survives when we neglect the satellites, and (2·39) becomes

$$\Delta_1\sigma_t^{T*}=-i\int_0^t dt_1\sum_s\langle F_{st}^{00}(0)\rangle_F\langle I_s^{0*}\rangle_s[I_t^0,\ \sigma_t^{T*}(0)]. \tag{2·41}$$

The second order term is given by

$$\begin{aligned}\Delta_2\sigma_t^{T*}=-\sum_{\lambda}\sum_{\lambda'}\sum_{s}\sum_{\mu_s}\sum_{\mu_{s'}}\sum_{\mu_t}\sum_{\mu_{t'}}\int_0^t dt_1\int_0^{t_1}dt_2\,e^{i(\omega_\lambda+\omega_{\mu_s}+\omega_{\mu_t}{}^0)t_1+i(\omega_{\lambda'}+\omega_{\mu_s}{}'^0+\omega_{\mu_t}{}'^0)t_2}\\
\times\{\langle F_{st}^{\bar{\mu}_s\bar{\mu}_t}(\lambda)F_{st}^{\bar{\mu}_{s'}\bar{\mu}_{t'}}(\lambda')\rangle_F\langle I^{\mu_s*}(t_1)\,I^{\mu_{s'}*}(t_2)\rangle_s I^{\mu_t}I^{\mu_{t'}}\sigma_t^{T*}(0)\\
+\langle F_{st}^{\bar{\mu}_{s'}\bar{\mu}_{t'}}(\lambda')F_{st}^{\bar{\mu}_s\bar{\mu}_t}(\lambda)\rangle_F\langle I^{\mu_{s'}*}(t_2)I^{\mu_s}(t_1)\rangle_s\sigma^{T*}(0)\,I^{\mu_{t'}}I^{\mu_t}\\
-\langle F_{st}^{\bar{\mu}_{s'}\bar{\mu}_{t'}}(\lambda')F_{st}^{\bar{\mu}_s\bar{\mu}_t}(\lambda)\rangle_F\langle I^{\mu_{s'}*}(t_2)I^{\mu_s*}(t_1)\rangle_s I^{\mu_t}\sigma_t^{T*}(0)\,I^{\mu_{t'}}\\
-\langle F_{st}^{\bar{\mu}_s\bar{\mu}_t}(\lambda)F_{st}^{\bar{\mu}_{s'}\bar{\mu}_{t'}}(\lambda')\rangle_F\langle I^{\mu_s*}(t_1)I^{\mu_{s'}*}(t_2)\rangle_s I^{\mu_{t'}}\sigma_t^{T*}(0)\,I^{\mu_t}\}\end{aligned} \tag{2·42}$$

where

$$\langle I^{\mu_s*}(t_1)I^{\mu_{s'}*}(t_2)\rangle_s\equiv tr\{\sigma_s^{T*}(t_1)I^{\mu_s*}(t_1)I^{\mu_{s'}*}(t_2)\} \tag{2·43}$$

stands for the correlation function of I^{μ_s} and $I^{\mu_{s'}}$ in the rotating frame. By factorizing the spin density in a self-consistent manner, we have already taken into account the main part of the effect of interplay of the local fields G, G_s, and G_t, even though we retain

terms only up to the second order. From the definition of trace we have $\omega_\lambda+\omega_{\lambda'}=0$, also we may choose $\omega_{\mu_t}+\omega_{\mu_t'}=0$ and $\omega_{\mu_s}+\omega_{\mu_s'}=0$ corresponding to the neglect of satellites. Then (2·42) is considerably simplified. After these manipulations we may reduce (2·38) to an approximate differential equation, in which the part of time change due to the interaction G is given by

$$\frac{d}{dt}\Big)_G \sigma_t^{T*} = \frac{d}{dt}\Delta_1\sigma_t^{T*} + \frac{d}{dt}\Delta_2\sigma_t^{T*}, \tag{2·44}$$

where

$$\frac{d}{dt}\Delta_1\sigma_t^{T*} = -i\sum_s \langle F_{st}^{00}(0)\rangle_F \langle I_s^{0*}\rangle_s [I_t^0,\ \sigma_t^{T*}], \tag{2·45}$$

and

$$\begin{aligned}\frac{d}{dt}\Delta_2\sigma_t^{T*} = -\sum_s\sum_{\mu_s}\sum_{\mu_t}\int_0^t d\tau\, e^{i(\omega_{\mu_s}+\omega_{\mu_t}{}^0)\tau}\\ \times[\ \langle F_{st}^{\bar\mu_s\bar\mu_t}(\tau)F_{st}^{\mu_s\mu_t}(0)\rangle\langle I^{\mu_s*}(\tau)I^{\bar\mu_s*}(0)\rangle_s\{I^{\mu_t}I^{\bar\mu_t}\sigma_t^{T*} - I^{\bar\mu_t}\sigma_t^{T*}I^{\mu_t}\}\\ + \langle F_{st}^{\mu_s\mu_t}(0)F_{st}^{\bar\mu_s\bar\mu_t}(\tau)\rangle\langle I^{\bar\mu_s*}(0)I^{\mu_s*}(\tau)\rangle_s\{\sigma_t^{T*}I^{\bar\mu_t}I^{\mu_t} - I^{\mu_t}\sigma_t^{T*}I^{\bar\mu_t}\}\ ];\\ \tau \equiv t_1 - t_2 .\end{aligned} \tag{2·46}$$

Here we have used the correlation function of the lattice part of G which is defined by

$$\begin{cases}\langle F_{st}^{\bar\mu_s\bar\mu_t}(\tau)F_{st}^{\mu_s\mu_t}(0)\rangle \equiv \sum_\lambda e^{i\omega_\lambda\tau}\langle F_{st}^{\bar\mu_s\bar\mu_t}(\lambda)F_{st}^{\mu_s\mu_t}(\bar\lambda)\rangle_F,\\ \langle F_{st}^{\mu_s\mu_t}(0)F_{st}^{\bar\mu_s\bar\mu_t}(\tau)\rangle = \sum_\lambda e^{i\omega_\lambda\tau}\langle F_{st}^{\mu_s\mu_t}(\bar\lambda)F_{st}^{\bar\mu_s\bar\mu_t}(\lambda)\rangle_F.\end{cases} \tag{2·47}$$

If we come back to the static frame at this stage by remembering the relation $\sigma_t^{T*}(t) = \sum_{\mu_t} e^{i\omega_{\mu_t}{}^0 t}\sigma_t(\mu_t)$, and derive from (2·45) a differential equation which is satisfied by the expectation value $\langle Q_t\rangle$ of a spin operator Q_t of the t-system, the time variation due to the interaction G is given by

$$\begin{aligned}\frac{d}{dt}\Big)_G\langle Q_t\rangle = -i\langle[Q_t,\ E_t]\rangle - i\sum_s\langle F_{st}^{00}(0)\rangle_F\langle I_s^{0*}\rangle_s\langle[Q_t,\ I_t^0]\rangle\\ -\sum_\lambda\sum_s\sum_{\mu_s}\sum_{\mu_t}\int_0^t d\tau\, e^{i(\omega_\lambda+\omega_{\mu_s}+\omega_{\mu_t})\tau}\\ \times\{\langle F_{st}^{\bar\mu_s\bar\mu_t}(\lambda)F_{st}^{\mu_s\mu_t}(\bar\lambda)\rangle_F\langle I^{\mu_s*}(\tau)I^{\bar\mu_s*}(0)\rangle_s\langle[Q_t,\ I^{\mu_t}]I^{\bar\mu_t}\rangle\\ -\langle F_{st}^{\mu_s\mu_t}(\bar\lambda)F_{st}^{\bar\mu_s\bar\mu_t}(\lambda)\rangle_F\langle I^{\bar\mu_s*}(0)I^{\mu_s*}(\tau)\rangle_s\langle\bar I^{\bar\mu_t}[Q_t,\ I^{\mu_t}]\rangle\},\end{aligned} \tag{2·48}$$

or by

$$\begin{aligned}\frac{d}{dt}\Big)_G\langle Q_t\rangle = -i\langle[Q_t,\ E_t]\rangle - i\sum_s\langle F_{st}^{00}(0)\rangle_F\langle I_s^{0*}\rangle_s\langle[Q_t,\ I_t^0]\rangle\\ -\sum_s\sum_{\mu_s}\sum_{\mu_t}\int_0^t d\tau\, e^{i(\omega_{\mu_s}+\omega_{\mu_t}{}^0)\tau}\end{aligned}$$

$$\times\{\langle F^{\bar{\mu}_s\bar{\mu}_t}_{st}(\tau)F^{\mu_s\mu_t}_{st}(0)\rangle_F\langle I^{\mu_s*}(\tau)I^{\bar{\mu}_s*}(0)\rangle_s\langle[Q_t,\ I^{\mu_t}]I^{\bar{\mu}_t}\rangle$$
$$-\langle F^{\mu_s\mu_t}_{st}(0)F^{\bar{\mu}_s\bar{\mu}_t}_{st}(\tau)\rangle_F\langle I^{\bar{\mu}_s*}(0)I^{\mu_s*}(\tau)\rangle_s\langle I^{\bar{\mu}_t}[Q_t,\ I^{\mu_t}]\rangle\}. \qquad (2\cdot49)$$

Essentially the same result is obtained for the case (b), i. e. $\theta_t\simeq\pi$, in Fig. 1. It should be noted that (2·49) involves two kinds of correlation function, for both of which we here assume smooth decay with finite lifetime, i. e.

$$\begin{cases}\sum_s\langle F^{\bar{\mu}_s\bar{\mu}_t}_{st}(\tau)F^{\mu_s\bar{\mu}_t}_{st}(0)\rangle\equiv\sum_s\langle F^{\bar{\mu}_s\bar{\mu}_t}_{st}F^{\mu_s\mu_t}_{st}\rangle e^{-\phi_{\mu_s\mu_t}\tau}\equiv\sigma^s_{t,\mu_s\mu_t}{}^2e^{-\phi_{\mu_s\mu_t}\tau}, & (2\cdot50)\\ \langle I^{\bar{\mu}_s*}(\tau)I^{\mu_s*}(0)\rangle_s\equiv\langle I^{\mu_s}I^{\bar{\mu}_s}\rangle_s f_{\mu_s}{}^*(\tau)\equiv\langle I^{\mu_s}I^{\bar{\mu}_s}\rangle_s e^{-\Phi_{\mu_s}{}^*\tau}, & (2\cdot51)\end{cases}$$

The latter is the correlation function of the operator I^{μ_s} in the rotating frame, and, as will be shown later, can be calculated from the equation describing the s-system so long as the Hartree approximation is valid. However, we have simply assumed its lifetime as $(\Phi_{\mu_s}{}^*)^{-1}$, so that we may obtain a clear qualitative idea. Let us introduce an abbreviation for the time integral which appears in (2·49)

$$\Phi^{\bar{\mu}_s\bar{\mu}_t*}\equiv\sum_s\int_0^t d\tau\, e^{i(\omega_{\mu_s}+\omega_{\mu_t}{}^0)\tau}\langle F^{\bar{\mu}_s\bar{\mu}_t}_{st}(\tau)F^{\mu_s\mu_t}_{st}(0)\rangle_F\langle I^{\bar{\mu}_s*}(\tau)I^{\bar{\mu}_s*}(0)\rangle_s/\langle I^{\mu_s}I^{\bar{\mu}_s}\rangle_s$$

$$=\sigma^s_{t,\mu_s\mu_t}{}^2\int_0^t d\tau\,\exp\{i(\omega_{\mu_s}+\omega_{\mu_t}{}^0)-(\phi_{\mu_s\mu_t}+\Phi_{\mu_s}{}^*)\}\tau$$

$$=\sigma^s_{t,\mu_s\mu_t}{}^2\left[\frac{\exp\{i(\omega_{\mu_s}+\omega_{\mu_t}{}^0)-(\phi_{\mu_s\mu_t}+\Phi_{\mu_s}{}^*)\}t-1}{i(\omega_{\mu_s}+\omega_{\mu_t}{}^0)-(\phi_{\mu_s\mu_t}+\Phi_{\mu_s}{}^*)}\right]. \qquad (2\cdot52)$$

Here we have five characteristic frequencies, i. e. ω_s, ω_t, $\sigma^s_{t,\mu s\mu t}$, $\phi_{\mu_s\mu_t}$, Φ_{μ_s}, and according to the relative magnitude of these frequencies we may classify several different cases. Let us remember the fact that $\Phi^{\mu_s\bar{\mu}_t*}$ is essentially the rate of phase variation of $\langle Q_t\rangle$, so that $\langle Q_t\rangle$ becomes practically vanishing as soon as the product $\Phi^{\mu\nu}t$ becomes greater than unity. Two typical cases are discriminated according to the asymptotic behaviour of $\Phi^{\mu\nu}$.

(A) The case in which $\mathrm{Max}(\omega_{\mu_s},\ \omega_{\mu_t}{}^0,\ \phi_{\mu_s\mu_t},\ \Phi_{\mu_s}{}^*)\gg\sigma^s_{t,\mu_s\mu_t}$.

$$\Phi^{\bar{\mu}_s\bar{\mu}_t*}\sim\Phi^{\bar{\mu}_s\bar{\mu}_t}\equiv\Phi'^{\bar{\mu}_s\bar{\mu}_t*}+i\Phi''^{\bar{\mu}_s\bar{\mu}_t*}\equiv\sigma^s_{t,\mu_s\mu_t}{}^2\int_0^\infty d\tau\, e^{\{i(\omega_{\mu_s}+\omega_{\mu_t}{}^0)-\phi_{\mu_s\mu_t}\}\tau}f_{\mu_s}{}^*(\tau)$$

$$\equiv\sigma^s_{t,\mu_s\mu_t}{}^2\left\{\frac{i(\omega_{\mu_s}+\omega_{\mu_t}{}^0)+(\phi_{\mu_s\mu_t}+\Phi_{\mu_s}{}^*)}{(\omega_{\mu_s}+\omega_{\mu_t}{}^0)^2+(\phi_{\mu_s\mu_t}+\Phi_{\mu_s}{}^*)^2}\right\} \qquad (2\cdot53)$$

(B) The case in which $\mathrm{Max}(\omega_{\mu_s},\ \omega_{\mu_t}{}^0,\ \phi_{\mu_s\mu_t},\ \Phi^{\mu_s*})\ll\sigma^s_{t,\mu_s\mu_t}$.

$$\Phi^{\bar{\mu}_s\bar{\mu}_t*}\sim\sigma^s_{t,\mu_s\mu_t}{}^2t. \qquad (2\cdot54)$$

Owing to the condition $\omega_s\gg\sigma^s_{t,\mu_s\mu_t}$ and $\omega_t\gg\sigma^s_{t,\mu_s\mu_t}$, which are valid except in the case of very weak constant field, the case $\mu_s\neq0$ or $\mu_t\neq0$ belongs to the situation (A) and $\Phi^{\bar{\mu}_s\bar{\mu}_t}$ is independent of time, whereas either situation (A) or (B) can exist for the case $\mu=0$. Up to this point we have discussed the equation of motion satisfied by the t-

system which is not saturated but traced by a weak radiation. In exactly the same manner we may derive the equation of motion satisfied by the s-system which is saturated by a strong resonant radiation. As a counterpart of (2·49), we are led to find

$$\left(\frac{d}{dt}\right)_G \langle Q_s^T \rangle = -i\langle [Q_s^T,\ E_s^{Te}] \rangle - i\sum_{\nu_s}\sum_{t} a_{0\nu_s}\langle F_{st}^{(0)}(0)\rangle_F \langle [Q_s^T,\ I(\nu_s)]\rangle$$

$$-\sum_{\mu_s}\sum_{\nu_s}\sum_{\nu_s'}\sum_{t}\sum_{\mu_t} a_{\mu_s\nu_s}a_{\bar{\mu}_s\nu_s'}\int_0^t d\tau\, e^{i(\omega_{\mu_s}-\omega_{\nu_s}'+\omega_{\mu_t})\tau}$$

$$\times\{\langle F_{st}^{\bar{\mu}_s\bar{\mu}_t}(\tau)F_{st}^{\mu_s\mu_t}(0)\rangle_F \langle I^{\mu_t *}(\tau)I^{\bar{\mu}_t *}(0)\rangle_t \langle [Q_s^T,\ I(\nu_s)]I(\nu_s')\rangle$$

$$-\langle F_{st}^{\mu_s\mu_t}(0)F^{\bar{\mu}_s\bar{\mu}_t}(\tau)\rangle_F \langle I^{\bar{\mu}_t *}(0)I^{\mu_t *}(\tau)\rangle_t \langle I(\nu_s')[Q_s^T,\ I(\nu_s)]\rangle\}. \quad (2\cdot 55)$$

Here, in contrast to (2·33), Q_s^T is a spin operator of the s-system in the rotating frame, and E_s^{Te} is due to the effective field acting on the s-system in the rotating frame. Accordingly, $I(\nu_s)$ is associated with ν_s quantum jump against this effective field, whereas $\langle I^{\mu_t *}(0)I^{\nu_t *}(\tau)\rangle_t$ is the correlation function of I^{μ_t} in the rotating frame and is written in a form

$$\langle I^{\bar{\mu}_t *}(0)I^{\mu_t *}(\tau)\rangle = e^{i\Delta_t t}f_{\mu_t}(\tau) = e^{i\Delta_t t-\Phi_{\mu_t}t}, \quad (2\cdot 56)$$

where $f_{\mu_t}(\tau)$ corresponds to the relaxation function in a frame which is rotating with the frequency ω_t, because $H_t \simeq 0$. Let us assume $f_{\mu_t}(\tau)$ has a relaxation time $\Phi_{\mu_t}^{-1}$, which can be calculated from the equation governing the motion of the t-system so long as the Hartree approximation is valid. We may also define

$$\Phi^{\bar{\mu}_s\bar{\nu}_s\bar{\mu}_t} \equiv \sum_t \int_0^t d\tau\, e^{i(\omega_{\mu_s}+\omega_{\nu_s}+\omega_{\mu_t}^0)\tau}\langle F_{st}^{\bar{\mu}_s\bar{\mu}_t}(\tau)F_{st}^{\mu_s\mu_t}(0)\rangle$$

$$\times\langle I^{\bar{\mu}_t *}(0)I^{\mu_t *}(\tau)\rangle_t \langle I^{\bar{\mu}_t *}I^{\mu_t *}\rangle_t$$

$$=\sigma_{s,\mu_s\mu t}^{t\,2}\int_0^t d\tau \exp\{i(\omega_{\mu_s}+\omega_{\nu_s}+\omega_{\mu_t}^0)-(\phi_{\mu_s\mu_t}+\Phi_{\mu_t})\}\tau$$

$$=\sigma_{s,\mu_s\mu_t}^{t\,2}\left\{\frac{\exp\{i(\omega_{\mu_s}+\omega_{\nu_s}+\omega_{\mu_t}^0)-(\phi_{\mu_s\mu_t}+\Phi_{\mu_t})\}t-1}{i(\omega_{\mu_s}+\omega_{\nu_s}+\omega_{\mu_t}^0)-(\phi_{\mu_s\mu_t}+\Phi_{\mu_t})}\right\}, \quad (2\cdot 57)$$

and again two extreme cases are discriminated with respect to the asymptotic behaviour of $\Phi^{\mu_s\bar{\nu}_s\nu_t}$.

(A) The case in which $\mathrm{Max}(\omega_{\mu_s},\ \omega_{\nu_s},\ \omega_{\mu_t}^0,\ \phi_{\mu_s\mu_t},\ \Phi_{\mu_t}) \gg \sigma_{s,\mu_s\mu_t}^t$.

$$\Phi^{\bar{\mu}_s\bar{\nu}_s\bar{\mu}_t} \equiv \Phi'^{\bar{\mu}_s\bar{\nu}_s\bar{\mu}_t} + i\Phi''^{\bar{\mu}_s\bar{\nu}_s\bar{\mu}_t} = \sigma_{s,\mu_s\mu_t}^{t\,2}\int_0^\infty d\tau\, e^{\{i(\omega_{\mu_s}+\omega_{\nu_s}+\omega_{\mu_t}^0)-\phi_{\mu_s\mu_t}\}\tau}f_{\mu_t}(\tau)$$

$$\sim \sigma_{s,\mu_s\mu_t}^{t\,2}\left\{\frac{i(\omega_{\mu_s}+\omega_{\nu_s}+\omega_{\mu_t}^0)+(\phi_{\mu_s\mu_t}+\Phi_{\mu_t})}{(\omega_{\mu_s}+\omega_{\nu_s}+\omega_{\mu_t}^0)^2+(\phi_{\mu_s\mu_t}+\Phi_{\mu_t})^2}\right\}. \quad (2\cdot 58)$$

(B) The case in which $\mathrm{Max}(\omega_{\mu_s},\ \omega_{\nu_s},\ \omega_{\mu_t}^0,\ \phi_{\mu_s\mu_t}) \ll \sigma_{s,\mu_s\mu_t}^t$.

$$\Phi^{\bar{\mu}_s\bar{\nu}_s\bar{\mu}_t} \sim \sigma'^{2}_{s,\mu_s\mu_t} t. \qquad (2\cdot59)$$

Owing to the condition $\omega_s \gg \sigma'_{s,\mu_s\mu_t}$ and $\omega_t \gg \sigma'_{s,\mu_s\mu_t}$, which is valid except in the case of very weak constant field, the case $\mu_s \neq 0$ or $\mu_t \neq 0$ belongs to the situation (A), and $\Phi^{\bar{\mu}_s\bar{\nu}_s\bar{\mu}_t}$ is practically independent of the saturation level ($\because \omega_s \gg \tilde{\omega}_s$)

$$\Phi^{\bar{\mu}_s\bar{\nu}_s\bar{\mu}_t} \equiv \Phi'^{\bar{\mu}_s\bar{\nu}_s\bar{\mu}_t} + i\Phi''^{\bar{\mu}_s\bar{\nu}_s\bar{\mu}_t}$$

$$\doteqdot \sigma'^{2}_{s,\mu_s\mu_t}\left\{\frac{(\phi_{\mu_s\mu_t}+\Phi_{\mu_t})+i(\omega_{\mu_s}+\omega_{\mu_t}{}^0)}{(\omega_{\mu_s}+\omega_{\mu_t}{}^0)^2+(\phi_{\mu_s\mu_t}+\Phi_{\mu_t})^2}\right\}. \qquad (2\cdot60)$$

Either situation (A) or (B) can exist for the case $\mu_s=0$ and $\mu_t=0$.

§ 3. The case of rapidly fluctuating lattice

——Overhauser polarization effect

When the lattice system F fluctuates more rapidly than the relative precession due to the distribution of the local field arising from the interaction G, that is the condition $\sigma_{00} \ll \phi_{00}$ holds, it belongs to the situation (A) and (2·53) and (2·57) can be applied. This means that we have a perfect definition of relaxation time in the sense of linear irreversible process. As there holds a condition $\sigma_{\mu_s\mu_t} \gtrsim \Phi_{\mu_s}{}^*$ in general, here we have the situation $\phi_{\mu_s\mu_t} \gg \Phi_{\mu_s}{}^*$. In other words, we may replace $\langle I^{\mu_s *}(\tau) I^{\bar{\mu}_s *}(0)\rangle_s$ by $\langle I^{\mu_s} I^{\bar{\mu}_s}\rangle_s$ in (2·49). In this case only the terms satisfying $\omega_\lambda + \omega_{\mu_s} + \omega_{\mu_t}{}^0 = 0$ survive and we have

$$\Phi^{\mu_s\mu_t *} = e^{-\beta\hbar(\omega_{\mu_s}+\omega_{\mu_t}{}^0)}\Phi^{\bar{\mu}_s\bar{\mu}_t *}. \qquad (3\cdot1)$$

By virtue of this relation, the equation (2·49) now becomes

$$\left(\frac{d}{dt}\right)_G \langle Q_t\rangle = -i\langle[Q_t,\ E_t]\rangle - i\sum_s \langle F^{00}_{st}(0)\rangle_F \langle I^{0*}_s\rangle_s \langle[Q_t,\ I^0_t]\rangle$$

$$-\sum_{\mu_s}\sum_{\mu_t}\Phi^{\bar{\mu}_s\bar{\mu}_t *}\{\langle I^{\mu_s} I^{\bar{\mu}_s}\rangle_s \langle[Q_t,\ I^{\mu_t}]I^{\bar{\mu}_t}\rangle - e^{-\beta\hbar(\omega_{\mu_s}+\omega_{\mu_t}{}^0)}\langle I^{\bar{\mu}_s} I^{\mu_s}\rangle_s \langle I^{\bar{\mu}_t}[Q_t,\ I^{\mu_t}]\rangle\}. \qquad (3\cdot2)$$

If the strong radiation is exactly resonant with the s-system, $\omega_{\mu_s} > \omega_{\mu_s}{}^0 = -\mu_s\omega_s$, and (3·2) essentially indicates the polarization effect which was first proposed by Overhauser. Actually, however, there also exists a relaxation due to the local field G_t, which was treated in paper I in detail. Let us assume that this latter relaxation process belongs to situation (A) and add the relevant terms by using the result in paper I. Remembering $a_{\mu\nu} = \delta_{\mu\nu}$ ($\because \theta_t \simeq 0$), we have the complete equation describing the t-system, which is being traced, as follows:

$$\frac{d}{dt}\langle Q_t\rangle = -i\langle[Q_t,\ E_t]\rangle - i\{\langle F^0_t(0)\rangle + \sum_s \langle I^0_s\rangle_s \langle F^{00}_{st}(0)\rangle_F\}\langle[Q_t,\ I^0_t]\rangle$$

$$-\sum_{\mu_t}\Phi^{\bar{\mu}_t}\{\langle[Q_t,\ I^{\mu_t}]I^{\bar{\mu}_t}\rangle - e^{-\beta\hbar\omega_{\mu_t}{}^0}\langle I^{\bar{\mu}_t}[Q_t,\ I^{\mu_t}]\rangle\}$$

$$-\sum_{\mu_s}\sum_{\mu_t}\Phi^{\bar{\mu}_s\bar{\mu}_t*}\{\langle I^{\mu_s}I^{\bar{\mu}_s}\rangle_s\langle[Q_t,\ I^{\mu_t}]I^{\bar{\mu}_t}\rangle-e^{-\beta\hbar(\omega_{\mu_s}{}^0+\omega_{\mu_t}{}^0)}\langle I^{\bar{\mu}_s}I^{\mu_s}\rangle_s\langle I^{\bar{\mu}_t}[Q_t,\ I^{\mu_t}]\rangle\} \tag{3·3}$$

(in the static system) .

Exactly similar manipulation of (2·55) including the local field relaxation due to G_t yields the following complete equation describing the s-system which is being saturated.

$$\frac{d}{dt}\langle Q_s^T\rangle=-i\langle[Q_s^T,\ E_s^{Te}]\rangle-i\sum_{\nu_s}a_{0\nu_s}\{\langle F_s^0(0)\rangle+\sum_t\langle I_t^0\rangle_t\langle F_{st}^{00}(0)\rangle_F\}\langle[Q_s^T,\ I(\nu_s)]$$

$$-\sum_{\mu_s}\sum_{\nu_s}\sum_{\nu_s'}a_{\mu_s\nu_s}a_{\bar{\mu}_s\bar{\nu}_s'}\Phi^{\bar{\mu}_s\nu_s'}\{\langle[Q_s^T,\ I(\nu_s)]I(\nu_s')\rangle-e^{-\beta\hbar(\omega_{\mu_s}+\omega_{\bar{\nu}_s}')}\langle I(\nu_s')[Q_s^T,I(\nu_s)]\rangle\}$$

$$-\sum_{\mu_s}\sum_{\nu_s}\sum_{\nu_s'}\sum_{\mu_t}a^{\bar{\mu}_s\nu_s}a^{\bar{\mu}_s\nu_s'}\Phi^{\bar{\nu}_s\nu_s'\bar{\mu}_t}$$

$$\times\{\langle I^{\mu_t}I^{\bar{\mu}_t}\rangle_t\langle[Q_s^T,\ I(\nu_s)]I(\nu_s')\rangle-e^{-\beta\hbar(\omega_{\mu_s}+\omega_{\bar{\nu}_s}'+\omega_{\mu_t}{}^0)}\langle I^{\bar{\mu}_t}I^{\mu_t}\rangle_t\langle I(\nu_s')[Q_s^T,\ I(\nu_s)]\rangle\}. \tag{3·4}$$

(in the rotating system)

It should be noted that (3·4) is valid in the rotating frame in contrast to (3·3) which is valid in the static frame.

§ 3. A. *Coupled free induction*

In the special case in which there is no strong radiation field resonant with the s-system, we have $\theta_s=0$, $a_{\mu_s\nu_s}=\delta_{\mu_s\nu_s}$, and $\mu_s=\nu_s=-\nu_s'$, so that (3·4) becomes

$$\frac{d}{dt}\langle Q_s\rangle=-i\langle[Q_s,\ E_s]\rangle-i\{\langle F_s^0(0)\rangle+\sum_t\langle I_t^0\rangle_t\langle F_{st}^{00}(0)\rangle_F\}\langle[Q_s,\ I_s^0]\rangle$$

$$-\sum_{\mu_s}\Phi^{\bar{\mu}_s}\{\langle[Q_s,\ I^{\mu_s}]I^{\bar{\mu}_s}\rangle-e^{-\beta\hbar\omega_{\mu_s}{}^0}\langle I^{\mu_s}[Q_s,\ I^{\mu_s}]\rangle\}$$

$$-\sum_{\mu_t}\sum_{\mu_s}\Phi^{\bar{\mu}_s\mu_t}\{\langle I^{\bar{\mu}_t}I^{\mu_t}\rangle_t\langle[Q_s,\ I^{\mu_s}]I^{\bar{\mu}_s}\rangle-e^{-\beta\hbar(\omega_{\mu_s}{}^0+\omega_{\mu_t}{}^0)}\langle I^{\bar{\mu}_t}I^{\mu_t}\rangle_t\langle I^{\mu_s}[Q_s,\ I^{\mu_s}]\rangle\} \tag{3·5}$$

(in the static system),

and in (3·3) $\Phi^{\mu_s\mu_t*}$ becomes $\Phi^{\mu_s\mu_t}$. Therefore, equation (3·3) and (3·5) are completely analogous to each other and form a set of simultaneous equations governing the coupled free induction of two kinds of spin system. However, in these equations the last term on the right-hand side involves quadratic terms in spin component and it is not easy to solve the equations for a general value of spin I_s or I_t. The only exception is the case $I_s=I_t=\frac{1}{2}$, in which the above set of equations becomes exactly linear in spin components as is shown in Appendix A. On the other hand, we have for the magnetic resonance in particular a condition

$$\hbar\omega_s\ll kT \quad\text{and}\quad \hbar\omega_t\ll kT \tag{3·6}$$

except, perhaps, in the lowest temperature range. (3·6) is often called the low frequency

condition. In this situation the equation (3·3) can be approximated by

$$\frac{d}{dt}\langle Q_t\rangle = -i\langle[Q_t,\ E_t]\rangle - i\{\langle F_t^0(0)\rangle + \sum_s \langle I_s^0\rangle_s \langle F_{st}^{00}(0)\rangle_F\}\langle[Q_t,\ I_t^0]\rangle$$
$$- \sum_{\mu_t} \Phi^{\bar{\mu}_t}\{\langle[[Q_t,\ I^{\mu_t}]I^{\bar{\mu}_t}]\rangle + \beta\hbar\,\omega_{\mu_t}{}^0\langle I^{\bar{\mu}_t}[Q_t,\ I^{\mu_t}]\rangle_0\}$$
$$- \sum_{\mu_s}\sum_{\mu_t} \Phi^{\bar{\mu}_s\bar{\mu}_t *}\{\langle I^{\mu_s}I^{\bar{\mu}_s}\rangle_s\langle[[Q_t,\ I^{\mu_t}]I^{\bar{\mu}_t}]\rangle + \langle[I^{\mu_s}I^{\bar{\mu}_s}]\rangle_S\langle I^{\bar{\mu}_t}[Q_t,\ I^{\mu_t}]\rangle$$
$$+ \beta\hbar\,(\omega_{\mu_s}{}^0 + \omega_{\mu_t}{}^0)\langle I^{\bar{\mu}_s}I^{\mu_s}\rangle_s\langle I^{\bar{\mu}_t}[Q_t,\ I^{\mu_t}]\rangle_0\}, \quad (3\cdot7)$$

where we have retained the terms only up to the first order in $(\hbar\omega_s/kT)$ or $(\hbar\omega_t/kT)$, and $\langle\ \rangle_0$ indicates the simple trace operation replacing the density matrix by unity. Exactly similar reduction is valid also for (3·4). Thus a set of *linear* simultaneous differential equations has been obtained. For the special case of coupled free induction the simultaneous equations are nothing but the equations proposed by Solomon[10] to describe the transient Overhauser effect in liquid H^1F^{19}. The same type of equations have also been proposed by Hasegawa[17] in analysing the coupled *s*-electron and *d*-electron system in the dilute alloy problem.

§ 3. B. *Overhauser effect*

In order to examine the effect of saturation of the *s*-system we must first solve (3·4). The treatment is, however, largely parallel to that which was presented in § 2 of paper I, and under the low frequency condition (3·6), the equation (3·4) becomes

$$\frac{d}{dt}\begin{bmatrix}\langle I_{sx}\rangle\\ \langle I_{sy}\rangle\\ \langle I_{sz}\rangle\end{bmatrix} + \begin{bmatrix}\Phi_{sx} & -\Delta_s+\Phi_{sxy} & \Phi_{sxz}\\ \Delta_s+\Phi_{syx} & \Phi_{sy} & -\omega_{1s}+\Phi_{syz}\\ \Phi_{szx} & \omega_{1s}+\Phi_{szy} & \Phi_{sz}\end{bmatrix}\begin{bmatrix}\langle I_{sx}\rangle\\ \langle I_{sy}\rangle\\ \langle I_{sz}\rangle\end{bmatrix} - \begin{bmatrix}\Phi'_{sx}\\ \Phi'_{sy}\\ \Phi'_{sz}\end{bmatrix}\langle I_t^0\rangle$$
$$= \frac{\hbar I_s(I_s+1)}{3kT}\begin{bmatrix}\Phi^*_{sx} & \tilde{\Phi}_{sx} & -\Phi'^t_{sx}\\ \Phi^*_{sy} & \tilde{\Phi}_{sy} & -\Phi'^t_{sy}\\ \Phi^*_{sz} & \tilde{\Phi}_{sz} & -\Phi'^t_{sz}\end{bmatrix}\begin{bmatrix}\omega\\ \tilde{\omega}_s\\ \omega_t\end{bmatrix} \quad (3\cdot8)$$

in the rotating system of coordinates. Note that the *z*-axis is along the direction of constant field in this equation. The relaxation constants Φ, Φ^*, $\tilde{\Phi}$, Φ' and Φ'' can be expressed as linear combinations of the quantities $\Phi^{\mu_s\nu_s\mu_t}$, $\Phi_s^{\mu_s\nu_s}$, $\Phi^{\mu_s\nu_s,s}$, and $\Phi^{\mu_s\nu_s,t}$, and the actual coefficients are given in Table 1. Here $\Phi_s^{\mu_s\nu_s}$ are the relaxation constants for the *s*-system when the interaction G is virtually removed, and were defined in paper I as $\Phi^{\mu\nu}$. $\Phi^{\mu_s\nu_s\mu_t}$ was defined by (2·57), $\Phi^{\mu_s\nu_s,s}$ and $\Phi^{\mu_s\nu_s,t}$ are defined by

$$\Phi^{\mu_s\nu_s,s} \equiv \frac{I_t(I_t+1)}{3}(\Phi^{\mu_s\nu_s 0} + 2\Phi^{\mu_s\nu_s\bar{1}} + 2\Phi^{\mu_s\nu_s 1}), \quad (3\cdot9)$$

$$\Phi^{\mu_s\nu_s,t} \equiv \frac{2I_s(I_s+1)}{3}(\Phi^{\mu_s\nu_s\bar{1}} - \Phi^{\mu_s\nu_s 1}). \quad (3\cdot10)$$

Table 1. Relaxation constants (the quantities in the first row) expressed as linear combinations of $\Phi^{\mu\nu}$. The coefficients are tabulated using the abbreviations $c \equiv \cos\theta$ and $s \equiv \sin\theta$, where θ is the angle between z- and z_s'-axis. The bar on the top indicates a minus sign. Note that $\Phi'^{\mu\nu,t} = -\Phi'^{\nu\mu,t}$.

	$\Phi_s^{00}+\Phi^{00,s}$	$\Phi_s'^{01}+\Phi'^{01,s}$	$\Phi_s'^{10}+\Phi'^{10,s}$	$\Phi_s'^{11}+\Phi'^{11,s}$	$\Phi_s'^{1\bar{1}}+\Phi'^{1\bar{1}},s$	
Φ_{sx}	c^2	s^2		$1+c$	$1-c$	
Φ_{sy}	c^2	s^2	$2s^2$	$c(c+1)$	$c(c-1)$	
Φ_{sz}			$2s^2$	$(c+1)^2$	$(c-1)^2$	
Φ_{sx}^*			$\overline{2cs}$	$s(c+1)$	$s(c-1)$	
$\tilde{\Phi}_{sx}$		s		$s(c+1)$	$\overline{s(c-1)}$	
Φ_{sz}^*			$2\bar{s}^2$	$(c+1)^2$	$(c-1)^2$	
$\tilde{\Phi}_{sz}$				$(c+1)^2$	$\overline{(c-1)^2}$	
Φ_{sxz}	$\overline{cs}$	cs				
Φ_{szx}			$2cs$	$s(c+1)$	$s(c-1)$	

		$\Phi_s''^{01}+\Phi_s''^{01},s$	$\Phi_s''^{10}+\Phi''^{10},s$	$\Phi_s''^{11}+\Phi''^{11},s$	$\Phi_s''^{1\bar{1}}+\Phi''^{1\bar{1}},s$	$-\sum_t \langle I_t^0 \rangle_t \langle F_{st}^{00}(0) \rangle_F$
Φ_{sy}^*			$2\overline{cs}$	$s(c+1)$	$s(c-1)$	
$\tilde{\Phi}_{sy}$		$\bar{cs}$		$s(c+1)$	$\overline{s(c-1)}$	
Φ_{sxy}			$\overline{2s^2}$	$\overline{c(c+1)}$	$\overline{c(c-1)}$	-1
Φ_{syx}				$(c+1)$	$c-1$	-1
Φ_{syz}		s				
Φ_{szy}			$2cs$	$s(c+1)$	$s(c-1)$	

	$\Phi^{00,t}$	$\Phi'^{01,t}$	$\Phi'^{\bar{1}0,t}$	$\Phi'^{11,t}$	$\Phi'^{\bar{1}1,t}$	
$\Phi_{sx}{}^t$	$c^2\bar{s}$	s	$2c^2s$	$s(c+1)$	$\overline{s(c-1)}$	$-\frac{I_s(I_s+1)}{I_t(I_t+1)}\Phi_{sx}{}'^t$
$\Phi_{sy}{}^t$	cs	cs	$2cs$	$s(c+1)$	$s(c-1)$	$=\frac{I_s(I_s+1)}{I_t(I_t+1)}\Phi_{sy}{}'^t$
$\Phi_{sz}{}^t$	cs^2		$\overline{2s^2}$	$(c+1)^2$	$\overline{(c-1)^2}$	$=\frac{I_s(I_s+1)}{I_t(I_t+1)}\Phi_{sz}{}'^t$

Except in the case of very low value of the constant field $\Phi''' \ll \Phi''$ holds for $\omega \neq 0$, and the results are somewhat simplified as given in Table 2. In particular the frequency shift is given by

$$\Delta\omega_s = \langle F_s^0(0) \rangle + \sum_t \langle I_t^0 \rangle_t \langle F_{st}^{00}(0) \rangle_F + 2(\Phi_s''^1 + \Phi''^{1,s}) \qquad (3\cdot11)$$

in this case, and we note it contains a term which is proportional to the expectation value of I_{tz}. Φ'''s are seen to be related to the frequency shift; however, it is usually higher in order than Φ' and their contribution may usually be neglected. Upon this simplification the equation $(3\cdot8)$ is reduced to

$$\frac{d}{dt}\begin{bmatrix}\langle I_{sx}\rangle\\ \langle I_{sy}\rangle\\ \langle I_{sz}\rangle\end{bmatrix} + \begin{bmatrix}\Phi_{sx} & -\Delta_s & 0\\ \Delta_s & \Phi_{sy} & -\omega_{1s}\\ 0 & \omega_{1s} & \Phi_{sz}\end{bmatrix}\begin{bmatrix}\langle I_{sx}\rangle\\ \langle I_{sy}\rangle\\ \langle I_{sz}\rangle\end{bmatrix} - \begin{bmatrix}\Phi_{sx}'^t\\ \Phi_{sy}'^t\\ \Phi_{sz}'^t\end{bmatrix}\langle I_t^0\rangle$$

$$= \frac{\hbar I_s(I_s+1)}{3kT} \begin{bmatrix} \Phi_{sx}\omega_{1s} - \Phi'^{t}_{sx}\omega_t \\ -\Phi'^{t}_{sy}\omega_t \\ \Phi_{sz}\omega_s - \Phi'^{t}_{sz}\omega_t \end{bmatrix}. \qquad (3\cdot12)$$

At exact resonance, we have

$$\begin{cases} \Phi^{t}_{sx} = \Phi'^{01,t} + \Phi'^{11,t} + \Phi'^{1\bar{1},t} = \Phi'^{01,t} & (3\cdot13) \\ \Phi^{t}_{sy} = 0, & (3\cdot14) \\ \Phi^{t}_{sz} = 2\Phi'^{10,t} + \Phi'^{11,t} - \Phi'^{1\bar{1},t} = 4\Phi'^{11,t} & (3\cdot15) \end{cases}$$

by remembering the relation $\Phi^{\mu_s\nu_s,t} = -\Phi^{\bar{\mu}_s\nu_s,t}$ (cf. (3·10)). If we know the value of $\langle I_t^0 \rangle_t$, then it is possible to solve (3·12) to see the behaviour of the s-system explicitly. To describe the steady state solution it is convenient to define the saturation factor S by the following formula,

Table 2. Relaxation constants (the quantities in the first row) expressed as linear combinations of $\Phi^{\mu\nu}$, using the same notations as in Table 1. (The case in which $\Phi^{10}=\Phi^{11}=\Phi^{1\bar{1}}\equiv\Phi'$.)

	$\Phi_s^{00}+\Phi^{00,s}$	$\Phi_s'^{01}+\Phi'^{01,s}$	$\Phi_s'^{1}+\Phi'^{1,s}$	
Φ_{sx}	c^2	s^2	2	
Φ_{sy}	c^2	s^2	2	
Φ_{sz}			4	
Φ_{sx}^*				
$\tilde{\Phi}_{sx}$		s	2s	
Φ_{sz}^*			4	
$\tilde{\Phi}_{sz}$			4c	
Φ_{sxz}	cs	cs		
Φ_{szx}				

		$\Phi_s''^{01}+\Phi''^{01,s}$	$\Phi_s''^{1}+\Phi''^{1,s}$	$\langle F_s^0(0)\rangle + \sum_t \langle I_t^0\rangle \langle F^{00}(0)\rangle_F$
Φ_{sy}^*				
$\tilde{\Phi}_{sy}$		cs	2cs	
Φ_{sxy}			-2	-1
Φ_{syx}			-2	-1
Φ_{syz}		$\bar{s}$		
Φ_{szy}				

	$\Phi^{00,t}$	$\Phi'^{01,t}$	$\Phi'^{1,t}$	
Φ_{sx}^t	$\overline{c^2 s}$	s	$2s(c^2+1)$	$=\frac{I_s(I_s+1)}{I_t(I_t+1)}\Phi_{sx}'^{t}$
Φ_{sy}^t	$\overline{cs}$	cs		$=\frac{I_s(I_s+1)}{I_t(I_t+1)}\Phi_{sy}'^{t}$
Φ_{sz}^t	cs^2		4	$=\frac{I_s(I_s+1)}{I_t(I_t+1)}\Phi_{sz}'^{t}$

$$\langle I_s^0 \rangle \equiv \langle I_{sz} \rangle_e (1-S), \tag{3·16}$$

where

$$\langle I_{sz} \rangle_e \equiv \frac{\hbar\omega_s}{3kT} I_s(I_s+1). \tag{3·17}$$

In order to pursue the behaviour of $\langle I_t^0 \rangle$, we should solve (3·7) which is explicitly written as

$$\frac{d}{dt}\begin{bmatrix} \langle I_{tx} \rangle \\ \langle I_{ty} \rangle \\ \langle I_{tz} \rangle \end{bmatrix} + \begin{bmatrix} \Phi_{tx} & -(\omega_t + \Delta\omega_t) & 0 \\ \omega_t + \Delta\omega_t & \Phi_{ty} & 0 \\ 0 & 0 & \Phi_{tz} \end{bmatrix} \begin{bmatrix} \langle I_{tx} \rangle \\ \langle I_{ty} \rangle \\ \langle I_{tz} \rangle - \langle I_{tz} \rangle_e \end{bmatrix}$$

$$- \begin{bmatrix} 0 \\ 0 \\ \Phi_{tz}^s \end{bmatrix} (\langle I_{sz} \rangle - \langle I_{sz} \rangle_e) = 0. \tag{3·18}$$

Here the relaxation constants are expressed by a linear combination of $\Phi_t^{\mu_t \nu_t}$ and $\Phi^{\mu_s \mu_t}$, defined by (2·53), as follows.

$$\begin{cases} \Phi_{tx} = \Phi_{ty} = \Phi_t^{00} + 2\Phi_t'^{11} + \frac{1}{5} I_s(I_s+1) \{\Phi^{00} + 4\Phi'^{10} + 2\Phi'^{01} + 4\Phi'^{11} + 4\Phi'^{1\bar{1}}\} & (3·19) \\ \Phi_{tz} = 4\Phi_t'^{11} + \frac{4}{5} I_s(I_s+1) \{\Phi'^{01} + 2\Phi'^{11} + 2\Phi'^{1\bar{1}}\} & (3·20) \\ \Phi_{tz}^s = \frac{8}{5} I_s(I_s+1) \{\Phi'^{1\bar{1}} - \Phi'^{11}\}. & (3·21) \end{cases}$$

It should be noted that all the relaxation constants in (3·18) are functions of $\sigma_{\mu_s \mu_t}$ and $\phi_{\mu_s \mu_t}$ only, and do not depend on the degree of saturation of the s-system, owing to the situation $\sigma_{00} \gg \phi_{00}$. The shift $\Delta\omega_t$ in resonance frequency of the t-system is given by

$$\Delta\omega_t = \langle F_t^0(0) \rangle + \sum_s \langle I_s^0 \rangle_s \langle F_{st}^{00}(0) \rangle_F + 2\Phi_t'^{11} + \frac{2}{5} I_s(I_s+1) \{\Phi'^{0\bar{1}} + 4\Phi'^{1\bar{1}} + 4\Phi'^{\bar{1}\bar{1}}\}. \tag{3·22}$$

The second term on the right-hand side corresponds to the *Knight shift* when there is no effect of saturation, i.e. $\langle I_s^0 \rangle_s = \langle I_s^0 \rangle_e$. However, when the s-system is saturated by a resonant radiation, (3·22) indicates that the Knight shift is diminished depending on the degree of saturation. The largest effect of saturation of the s-system appears in the z-polarization of the t-system. To look into the explicit dependence let us solve (3·12) and (3·18) for a steady state condition. If we concentrate our attention on the z-components of spin vectors, we have

$$\begin{cases} \Phi_{sz}(1+\lambda) \langle I_{sz} \rangle_0 - \Phi_{sz}^t \langle I_{tz} \rangle = \Phi_{sz} \langle I_{sz} \rangle_e - \Phi_{sz}^t \langle I_{tz} \rangle_e, & (3·23) \\ -\Phi_{tz}^s \langle I_{sz} \rangle_0 + \Phi_{tz} \langle I_{tz} \rangle_0 = -\Phi_{tz}^s \langle I_{sz} \rangle_e + \Phi_{tz} \langle I_{tz} \rangle_e, & (3·24) \end{cases}$$

from (3·12) and (3·18) respectively. Here we used an abbreviation

$$\lambda \equiv \omega_{1s}^2 / \Phi_{sy}\,\Phi_{sz} \quad \text{(saturation parameter)} \tag{3·25}$$

and $\langle\ \ \rangle_0$ denotes the steady state value. Solving these equations, we have the steady state z-polarizations in terms of equilibrium polarizations and the saturation parameter λ, or of saturation factor S defined in (3·16).

$$\begin{cases} \langle I_{sz}\rangle_0 = \langle I_{sz}\rangle_e \left\{1 - \dfrac{\lambda \Phi_{sz}\,\Phi_{tz}}{(1+\lambda)\,\Phi_{sz}\,\Phi_{tz} - \Phi_{sz}^t\,\Phi_{tz}^s}\right\} \equiv \langle I_{sz}\rangle_e (1-S)\ , \\ \langle I_{tz}\rangle_0 = \langle I_{tz}\rangle_e \left\{1 - \dfrac{\lambda \Phi_{sz}\,\Phi_{tz}^s}{(1+\lambda)\,\Phi_{sz}\,\Phi_{tz} - \Phi_{sz}^t\,\Phi_{tz}^s}\left(\dfrac{\langle I_{sz}\rangle_e}{\langle I_{tz}\rangle_e}\right)\right\} \end{cases} \tag{3·26}$$

$$\begin{aligned} &= \langle I_{tz}\rangle_e \left\{1 - \left(\frac{\Phi_{tz}^s}{\Phi_{tz}}\right)\left(\frac{\langle I_{sz}\rangle_e}{\langle I_{tz}\rangle_e}\right) S\right\} \\ &= \langle I_{tz}\rangle_e \left\{1 + \frac{2\,(\Phi^{\prime 11} - \Phi^{\prime 1\bar{1}})}{\Phi_{tz}}\left(\frac{\omega_s}{\omega_t}\right)\frac{4I_s(I_s+1)}{3}\,S\right\}\ . \end{aligned} \tag{3·27}$$

The equation indicates clearly that the z-polarization of the t-system is affected when the s-system is saturated. This effect is most remarkable for the case in which $\omega_s \gg \omega_t$, and (3·27) is a generalized formula for the *Overhauser polarization effect* originally predicted for the electron-nucleus spin system.[1] Suppose the s-system has spin $I_s = \frac{1}{2}$ and the two systems are coupled through scalar product interaction, which is in turn the only mechanism for the lattice relaxation of the t-system, then we have

$$\lim_{S\to 1} \langle I_{tz}\rangle_0 = \frac{\hbar\,(\omega_t - \omega_s)}{3kT}\, I_t(I_t+1) \tag{3·28}$$

in the limit of extreme saturation. The situation may be described by saying that there exists a canonical distribution in the rotating frame[1(ii)] in this limit. However, this simple description is valid only in the special case defined above, and in general there can be no such simple description. Nevertheless, essentially the same effect is expected because $\Phi^{\prime 1\bar{1}} \neq \Phi^{\prime 11}$ for a general type of interaction. At the same time (3·11) indicates that there appears a shift in the resonance frequency of the s-system owing to the extra-polarization of the t-system. This is the so-called *Overhauser shift*[5] and was originally proposed[7] as a means of detecting the polarization effect. For the special case in which $I_s = I_t = \frac{1}{2}$, we need no low frequency condition to have a linear form of the differential equation. The treatment of this case is included in Appendix A.

§ 4. The case of rigid lattice

——*Saturational narrowing*

When the lattice system F does not actually fluctuate over many periods of the relative precession due to the distribution of the local field arising from the interaction G, i. e. the condition $\sigma_{00} \gg \phi_{00}$ holds, there is no assurance in general that we can draw any substantial information about the time variation of density matrix from its short time behaviour. However, it does yield valuable information in a special case in which the origin of the

width is due to many perturbers which have regular static arrangement. A Gaussian shape of distribution is expected in this case as a result of the central limit theorem, therefore one can reproduce the whole line shape in the absence of saturation once he has the value of the adiabatic second moment. Being aware of this restriction, we here limit ourselves to this special case, which is not without experimental example. In this case only the terms corresponding to $\omega_s = 0$ survive, and we may assume $\omega_{s\mu_s} \sim 0$ in (2·50) and in the definition of $\Phi^{\mu_s\bar{\mu}_t *}$ and $\Phi^{\mu_s\nu_{s'}\bar{\mu}_t}$. Let us further assume that in the absence of the interaction G the relaxation processes due to the local field G and G_s are of type (A). Then the differential equation governing the motion of the t-system becomes

$$\frac{d}{dt}\langle Q_t\rangle = -i\langle[Q_t,\ E_t]\rangle - i\{\langle F_t^0(0)\rangle + \sum_{\mu_t}{}'\langle I_t^{\mu_t}\rangle\langle F_t^{\bar{\mu}_t}(0)\rangle\}\langle[Q_t,\ I_t^0]\rangle$$
$$- \sum_{\mu_t}{}'\Phi^{\mu_t}\{\langle Q_t,\ I^{\mu_t}|I^{\bar{\mu}_t}\rangle - e^{-\beta\hbar\omega_{\mu_t}}\langle I^{\mu_t}|Q_t,\ I^{\mu_t}\rangle\}$$
$$- \sum_{\mu_s}{}'\sum_{\mu_t}{}'\Phi^{\bar{\mu}_s\bar{\mu}_t *}\{\langle I^{\mu_s}I^{\bar{\mu}_s}\rangle_s\langle[Q_t,\ I^{\mu_t}]I^{\bar{\mu}_t}\rangle - \langle I^{\bar{\mu}_s}I^{\mu_s}\rangle_s\langle I^{\bar{\mu}_t}[Q_t,\ I^{\mu_t}]\rangle\}$$
$$\text{(in the static system)}, \qquad (4\cdot1)$$

where $\Phi^{\bar{\mu}_s\bar{\mu}_t *}$ is defined by (2·52) keeping $\omega_{s\mu_s} = 0$. In the same manner the equation governing the motion of the s-system is obtained as

$$\frac{d}{dt}\langle Q_s^T\rangle = -i\langle[Q_s,\ E_s^{Te}]\rangle - i\sum_{\nu_s}{}'a_{\nu_s}\{\langle F^0(0)\rangle + \sum_t{}'\langle I_t^{\mu_t}\rangle_t\langle F_t^{\bar{\mu}_t}(0)\rangle_t\}\langle Q_s^T,\ I(\nu_s)\rangle$$
$$- \sum_{\mu_s}{}'\sum_{\nu_s}{}'a_{\mu_s\nu_s}a_{\bar{\mu}_s\nu_{s'}}\Phi^{\bar{\mu}_s\nu_{s'}}\{\langle Q_s^T,\ I(\nu_s)\rangle I(\nu_{s'}) - e^{-\beta\hbar\omega_{\mu_s}-\omega_{\nu_{s'}}}\langle I(\nu_{s'})\rangle\langle Q_s^T,\ I(\nu_s)\rangle\}$$
$$- \sum_{\mu_s}{}'\sum_{\nu_s}{}'\sum_{\nu_{s'}}{}'\sum_{\mu_t}{}'a_{\bar{\mu}_s\nu_s}a_{\mu_s\nu_{s'}}\Phi^{\mu_s\nu_{s'}\bar{\mu}_t}$$
$$\times\{\langle I^{\mu_t}I^{\bar{\mu}_t}\rangle_t\langle[Q_s^T,\ I(\nu_s)]I(\nu_{s'})\rangle - \langle I^{\bar{\mu}_t}I^{\mu_t}\rangle_t\langle I(\nu_{s'})[Q_s^T,\ I(\nu_s)]\rangle\}, \quad (4\cdot2)$$
$$\text{(in the rotating system)}$$

where $\Phi^{\bar{\mu}_s\nu_{s'}\mu_t}$ is defined by (2·57) keeping $\omega_{s\mu_s} = 0$. Under the low frequency condition (3·6) the equation (4·2) is explicitly written as

$$\frac{d}{dt}\begin{bmatrix}\langle I_{sx}\rangle\\ \langle I_{sy}\rangle\\ \langle I_{sz}\rangle\end{bmatrix} + \begin{bmatrix}\Phi_{sx} & -\Delta_s+\Phi_{sxy} & \Phi_{sxz}\\ \Delta_s+\Phi_{syx} & \Phi_{sy} & -\omega_{1s}+\Phi_{syz}\\ \Phi_{szx} & \omega_{1s}+\Phi_{szy} & \Phi_{sz}\end{bmatrix}\begin{bmatrix}\langle I_{sx}\rangle\\ \langle I_{sy}\rangle\\ \langle I_{sz}\rangle\end{bmatrix} - \begin{bmatrix}\Phi_{sx}^t\\ \Phi_{sy}^t\\ \Phi_{sz}^t\end{bmatrix}\langle I_t^0\rangle$$
$$= -\frac{\hbar I_s(I_s+1)}{3kT}\begin{bmatrix}\Phi_{sx}^{*0} & \tilde{\Phi}_{sx}^0\\ \Phi_{sy}^{*0} & \tilde{\Phi}_{sy}^0\\ \Phi_{sz}^{*0} & \tilde{\Phi}_{sz}^0\end{bmatrix}\begin{bmatrix}\omega\\ \tilde{\omega}_s\end{bmatrix}, \quad \text{(in the rotating system)} \qquad (4\cdot3)$$

where the relaxation constants are expressed by linear combinations of $\Phi^{\mu_s\nu_s}$ and $\Phi^{\mu_s\nu_{s'}\mu_t}$, which is defined by (2·57). The coefficients of linear combinations are formally the same as given in Table 1, except that Φ_{sx}^{*0}, $\tilde{\Phi}_{sx}^0$, etc., correspond to Φ_{sx}^*, $\tilde{\Phi}_{sx}$, etc., with the contribution from G absent. When $\Phi^{\mu_s\nu_{s'}} = \Phi^{\mu_s}$, we do have the simplified relation

given in Table 2. When compared with the corresponding equations, the largest difference lies in the fact that $\Phi^{\mu_s \nu \mu_t}$ involves $f_{\mu_t}(\tau)$ explicitly in its definition. As we have discussed in the previous section, the terms corresponding to $\mu_s \neq 0$ or $\mu_t \neq 0$ belongs to the situation (A) ; however, their contribution to the relaxation constants is negligible as compared with that of the term having $\mu_s = \mu_t = 0$. In the absence of saturation the term corresponding to $\mu_s = \mu_t = 0$ belongs to the situation (B), and the shape of the absorption is given by

$$f_s(\omega) = \left(\frac{1}{2\pi\sigma_s^2}\right)^{1/2} e^{-\frac{(\omega-\omega_s)^2}{2\sigma_s^2}}, \tag{4·4}$$

which has a half-value width

$$\sigma'_{s,00} = 1{,}175\,\sigma_s = 1.175\left(\frac{I_s(I_s+1)}{3}\right)^{1/2}\sigma_{s,00}. \tag{4·5}$$

On the other hand it is clear from (2·57) that the situation must be changed into (B) when the intensity of the radio-frequency field, i. e. $\tilde{\omega}_s$, becomes large enough. In the last case we have

$$\Phi^{0\nu_s 0} \sim \sigma_{s,00}^2 \int_0^\infty d\tau\, e^{i\omega\nu_s\tau} f_0^t(\tau)\ . \tag{4·6}$$

In order to observe the qualitative behaviour in the entire range, we may replace the Gaussian limit by a correponding Lorentzian line which has an equivalent half-value width. An example of the desired interpolation between (4·5) and (4·6) is given by

$$\frac{1}{3} I_s(I_s+1)\,\Phi^{0\nu_s 0} = \sigma'_{s,00}\left[1 - \exp\left\{-\sigma''_{s,00}\int_0^\infty d\tau\, e^{i\omega\nu_s\tau} f_0^t(\tau)\right\}\right], \tag{4·7}$$

where $\sigma''_{s,00}$ is defined by the relation

$$\sigma_s^2 = \frac{1}{3} I_s(I_s+1)\,\sigma_{s,00}^2 = \sigma'_{s,00}\sigma''_{s,00}. \tag{4·8}$$

Owing to the fact that $\Psi' \gg \Psi''$, the second order shift may be neglected in (3·11). In view of the fact that $\Phi'^{0\nu_s,s} \gg \Phi'^{1,s}$ and $\Phi_s'^{1} \gg \Phi'^{1,s}$ (4·3) is simplified to be

$$\frac{d}{dt}\begin{bmatrix}\langle I_{sx}\rangle\\ \langle I_{sy}\rangle\\ \langle I_{sz}\rangle\end{bmatrix} + \begin{bmatrix}\Phi_{sx} & -\Delta_s & 0\\ \Delta_s & \Phi_{sy} & -\omega_{1s}\\ 0 & \omega_{1s} & \Phi_{sz}\end{bmatrix}\begin{bmatrix}\langle I_{sx}\rangle\\ \langle I_{sy}\rangle\\ \langle I_{sz}\rangle\end{bmatrix} - \begin{bmatrix}\Phi_{sx}^t\\ \Phi_{sy}^t\\ \Phi_{sz}^t\end{bmatrix}\langle I_t^0\rangle = \frac{\hbar I_s(I_s+1)}{3kT}\begin{bmatrix}\Phi_{sx}^0\omega_{1s}\\ 0\\ \Phi_{sz}^0\,\omega_s\end{bmatrix}$$

(in the rotating frame), (4·9)

where Φ_{sx}^0 and Φ_{sz}^0 stands for the corresponding Φ_{sx} and Φ_{sz} with the contribution of G absent. At exact resonance the formulae (3·13), (3·14), and (3·15) apply also here, and we see that Φ_{sz}^t can be neglected as compared with Φ_{sx}^t.

Under similar conditions the equation (4·1) is explicitly written as

$$\frac{d}{dt}\begin{bmatrix}\langle I_{tx}\rangle\\ \langle I_{ty}\rangle\\ \langle I_{tz}\rangle\end{bmatrix}+\begin{bmatrix}\Phi_{tx} & -(\omega_t+\Delta\omega_t) & 0\\ \omega_t+\Delta\omega_t & \Phi_{ty} & 0\\ 0 & 0 & \Phi_{tz}\end{bmatrix}\begin{bmatrix}\langle I_{tx}\rangle\\ \langle I_{ty}\rangle\\ \langle I_{tz}\rangle\end{bmatrix}-\begin{bmatrix}0\\ 0\\ \Phi^s_{tz}\langle I_s^0\rangle\end{bmatrix}$$

$$=\begin{bmatrix}0\\ 0\\ \Phi^0_{tz}\langle I_{tz}\rangle_e\end{bmatrix}, \tag{4·10}$$

in which the relaxation constants are expressed as linear combinations of $\Phi_t^{\mu_t\nu_t}$ and $\Phi^{\mu_s\mu_t*}$ as follows.

$$\Phi_{tx}=\Phi_{ty}=\Phi_t^{00}+2\Phi_t'^{11}+\tfrac{1}{3}I_s(I_s+1)\{\Phi^{00*}+4\Phi'^{10*}+2\Phi'^{01*}+4\Phi'^{11*}+4\Phi'^{1\bar{1}*}\} \tag{4·11}$$

$$\Phi_{tz}=4\Phi_t'^{11}+\tfrac{4}{3}I_s(I_s+1)\{\Phi'^{01*}+2\Phi'^{11*}+2\Phi'^{1\bar{1}*}\} \tag{4·12}$$

$$\Phi^0_{tz}=4\Phi_t'^{11} \tag{4·13}$$

$$\Phi^s_{tz}=\tfrac{8}{3}I_s(I_s+1)\{\Phi'^{1\bar{1}*}-\Phi'^{11*}\}>0. \tag{4·14}$$

It should be remembered that $\Phi_{tx}=\Phi^0_{tx}$ if the lattice relaxation of the t-system is determined by G_t rather than G, whereas $\Phi^0_{tz}=0$ if it is determined by G only. The shift $\Delta\omega_t$ of the resonance frequency is given by

$$\Delta\omega_t=\langle F_t^0(0)\rangle+\sum_s\langle I_s^0\rangle\langle F_{st}^{(0)}(0)\rangle_F+2\Phi_t'^{11}+\tfrac{2}{3}I_s(I_s+1)\{\Phi'^{01*}+4\Phi'^{11*}+4\Phi'^{1\bar{1}*}\}. \tag{4·15}$$

If we compare these formulae with (3·19) - (3·22), the only difference lies in the definition (2·52) of $\Phi'^{s\nu_t*}$ so far as it involves $f_{\mu_s}^{s*}(\tau)$. The terms corresponding to $\mu_s\neq 0$ or $\mu_t\neq 0$ belong to situation (A), but their contribution being small, they may be neglected. In the absence of saturation the term corresponding to $\mu_s=\mu_t=0$ gives rise to a *Gaussian shape of absorption* (situation (B)) having a half-value width

$$\sigma'_{t,00}=1.175\,\sigma_t=1.175\left(\frac{I_t(I_t+1)}{3}\right)^{1/2}\sigma_{t\,00}, \tag{4·16}$$

provided the width coming from process (A) is negligible. On the other hand *the situation must be changed into (A) and the line shape becomes Lorentzian when the s-system is saturated*, for the lifetime of $f_{\mu_s}^{s*}(\tau)$ becomes shorter when the radio-frequency field becomes stronger (nutational modulation). In this limit we have

$$\Phi^{00*}\sim\sigma^2_{t,00}\int_0^\infty d\tau\, f_0^{s*}(\tau). \tag{4·17}$$

To interpolate between the two extremes, (4·16) and (4·17), we adopt the same convention as in (4·7), i. e.

$$\tfrac{1}{3}I_s(I_s+1)\Phi^{00*}\fallingdotseq\sigma'_{t,00}\left[1-\exp\{-\sigma''_{t,00}\int_0^\infty d\tau\, f_0^{s*}(\tau)\}\right], \tag{4·18}$$

where $\sigma''_{t,00}$ is defined by the relation

$$\sigma_t^2 \equiv \tfrac{1}{3} I_s(I_s+1)\,\sigma^2_{t,00} \equiv \sigma'_{t,00}\,\sigma''_{t,00}. \tag{4·19}$$

§ 4. A. *Saturational narrowing*

From (4·11) it is clear that the transverse relaxation of the t-system depends on the saturation of the s-system mainly through $\Psi^{00}*$, for the other terms are either independent of saturation or rather small in magnitude as compared with $\Psi^{00}*$. Accordingly, the transition of the situation from (B) to (A) means that the width of t-systsm resonance is diminished upon saturating the s-system resonance, which should properly be called "*saturational narrowing*". Let us examine the transverse relaxation a little more in detail to see just at what level the effect of saturation comes in. This may conveniently be done in three steps.

i) For the case in which the longitudinal relaxation of the t-system is largely determined by G_t, Φ_{tz} is essentially equal to $1/T_1$ for the free decay. Therefore $f_0^t(\tau) = e^{-\Phi_{tz}\tau}$ is easily obtained experimentally.

ii) From Table 2, (3·9), and (3·10), we have

$$\Phi_{sy} \simeq \Phi'^{01}_s + \Phi'^{01,s} \simeq \Phi'^{01}_s + \Phi'^{01,0}\left\{\frac{I_s(I_t+1)}{3}\right\}, \tag{4·20}$$

where

$$\Phi'^{01,0} = \mathrm{Re}\,\sigma^2_{s,00}\int_0^t d\tau\; e^{i\omega_{1s}\tau} f_0^t(\tau) \tag{4·21}$$

from (4·5). If we adopt the interpolation formula (4·7), we have

$$\tfrac{1}{3} I_t(I_t+1)\,\Phi'^{01,0} = \sigma'_{s,00}\left\{1 - e^{-\frac{\sigma'_{s,00}\Phi_{tz}}{\Phi_{tz}^2+\omega_{1s}^2}}\right\}. \tag{4·22}$$

As for the contribution arising from G_s we have already discussed it in paper I and have

$$\Psi'^{01}_s = \sigma'_{s0}\left\{1 - e^{-\frac{\sigma''_{s0}\phi_{s0}}{\phi_{s0}^2+\omega_{1s}^2}}\right\}, \tag{4·23}$$

so that the total transverse relaxation is given by

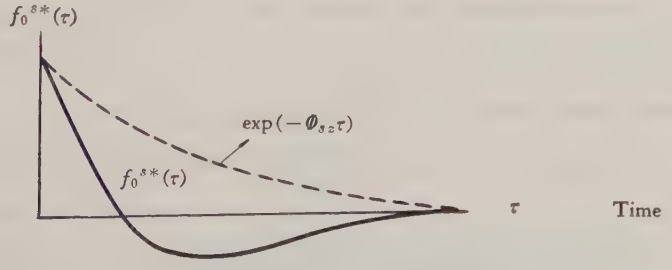


Fig. 2. The behaviour of $f_0^{s*}(\tau)$ as a function of time.

$$\Phi_{sy}=\sigma'_{s0}\left\{1-e^{-\frac{\sigma_{s0}''\phi_{s0}}{\phi_{s0}^2+\omega_{1s}^2}}\right\}+\sigma'_{s,00}\left\{1-e^{-\frac{\sigma_{s,00}''\Phi_{tz}}{\Phi_{tz}^2+\omega_{1s}^2}}\right\}. \tag{4·24}$$

iii) Having thus obtained the expression of Φ_{sy}, and neglecting the coupling due to G as small ($\because$ they have only terms corresponding to $\mu=0$), it is straightforward to calculate $f_0^{s*}(\tau)$ from (4·9). This is done in Appendix B and the result is given by (cf. (A·20))

$$f_0^{s*}(\tau)=\frac{1}{\Phi_2^- -\Phi_1^+}\left\{(\Phi_{sz}-\Phi_1^+)\,e^{-\Phi_2^- t}-(\Phi_{sz}-\Phi_2^-)\,e^{-\Phi_1^+ t}\right\}, \tag{4·25}$$

where (cf. (A·8))

$$\begin{pmatrix}\Phi_2^-\\ \Phi_1^+\end{pmatrix}=\frac{1}{2}\left\{(\Phi_{sy}+\Phi_{sz})\pm\sqrt{(\Phi_{sy}-\Phi_{sz})^2-4\omega_{1s}^2}\right\}. \tag{4·26}$$

The behaviour of (4·25) is shown in Fig. 2 and indicates clearly that the relaxation time of I_s^0 is decreased as a result of the forced oscillation by the radio-frequency field.

iv) Inserting (4·25) into (4·18) we have (cf. (A·23))

$$\frac{1}{3}I_s(I_s+1)\Phi^{00*}=\sigma'_{t,00}\left\{1-\exp\left(-\frac{\sigma''_{t,00}\Phi_{sy}}{\Phi_{sy}\Phi_{sz}+\omega_{1s}^2}\right)\right\}, \tag{4·27}$$

and according to the result of paper I

$$\Phi_t^{00}=\sigma'_{t0}\left\{1-\frac{1}{2}\exp\left(-\frac{\sigma''_{t0}}{\phi_{t0}}\right)\right\}\simeq\sigma'_{t0}. \tag{4·28}$$

Therefore the transverse relaxation of the t-system is approximately given by

$$\Phi_{ty}\simeq\Phi_t^{00}+\frac{1}{3}I_s(I_s+1)\Phi^{00*}=\sigma'_{t,0}+\sigma'_{t,00}\left\{1-e^{-\frac{\sigma_{t,00}''\Phi_{sy}}{\Phi_{sy}\Phi_{sz}+\omega_{1s}^2}}\right\}. \tag{4·29}$$

This is our general formula for the saturational narrowing, and indicates that the part $\sigma'_{t,00}$, due to the interaction G, is narrowed off when the s-system is saturated. Associated with this narrowing there appears an increase in the peak intensity of the absorption which is clearly seen in the steady state solution

$$\langle I_{ty}\rangle_0=\frac{\langle I_{tz}\rangle_e\Phi_{tz}^0+\langle I_{sz}\rangle_0\Phi_{tz}^s}{\Phi_{ty}\Phi_{tz}+\omega_{2t}^2}\cdot\omega_{2t}, \tag{4·30}$$

where $H_2\cos\omega_{2t}\,\gamma_t$ is the amplitude of a second weak radiation resonant with the t-system.

§ 4. B. *Depolarization effect*

As the steady state solution of (4·9) and (4·10) we have

$$\langle I_{sz}\rangle_0=\langle I_{sy}\rangle_0\Phi_{sy}/\omega_{1s}=\frac{\Phi_{sz}^0\Phi_{tz}\langle I_{sz}\rangle_e+\Phi_{tz}^0\Phi_{sz}^t\langle I_{tz}\rangle_e}{\Phi_{sz}\Phi_{tz}(1+\lambda)-\Phi_{tz}^s\Phi_{sz}^t}, \tag{4·31}$$

$$\langle I_{tz}\rangle_0=\frac{\Phi_{sz}\Phi_{tz}^0(1+\lambda)\langle I_{tz}\rangle_e+\Phi_{sz}^0\Phi_{tz}^s\langle I_{sz}\rangle_e}{\Phi_{sz}\Phi_{tz}(1+\lambda)-\Phi_{tz}^s\Phi_{sz}^t}, \tag{4·32}$$

where $\lambda = \omega_{1s}^{2} \Psi_{sy} \Psi_{sz}$. On saturating the s-system obviously $\langle I_{sz} \rangle_0 \to 0$, whereas $\langle I_{tz} \rangle_0 \to (\Psi_{tz}^{0} / \Psi_{tz}) \langle I_{tz} \rangle_e$. If an appreciable part of the lattice relaxation of the t-system arises from the s-t interaction, i. e. Ψ_{tz}^{0} is definitely smaller than Ψ_{tz}, this means that the limiting value of the z-polarization of the t-system is definitely smaller than $\langle I_{tz} \rangle_e$. In contrast to the situation discussed in the previous section we have a depolarization effect rather than a polarization. It should be borne in mind that in this section the saturational behaviour of the peak absorption and polarization are not really proportional to each other, for the ratio $(\Psi_{ty} / \omega_{2t})$ actually depends on the degree of saturation of the s-system.

§ 5. Comparison with experiment and discussion

§ 5. A. *Detection of Overhauser effect*

Let us summarize here various possible ways of detecting the polarization of the nuclear spin system (t-system) on saturating the electron spin (s-system) resonance.

(1) *Increase of resonance intensity*

As there is a relation $\langle I_{ty} \rangle_0 = \langle I_{tz} \rangle_0 (\omega_{2t} / \Psi_{ty})$, it is possible to observe the absorption as a direct indicator of z-polarization. This is the method used by Carver and Slichter[8] in their first confirmation of the Overhauser effect. We note that the direct proportionality is assured only in the case of rapidly fluctuating lattice, and not in the case of rigid lattice in general.

(2) *Overhauser shift*

As was mentioned in § 4, (3·11) indicates that the electron spin resonance frequency should change as a result of nuclear polarization. Kaplan[7] did this type of experiment, but his result does not seem to be so clear as that of Carver and Slichter.[8]

(3) A nuclear indicator of the electron spin depolarization which might be called *differential Knight shift.*

The Knight shift, which is proportional to the z-polarization of the s-system as shown in (3·22), should decline to nought on saturating the electron spin resonance. This effect could be an indicator of the depolarization of electron spins, although the author has never seen this effect reported. In principle this effect may be used to determine the absolute value of Knight shift. The point is that here we have no ambiguity in the choice of reference frequency with respect to which the Knight shift is measured. For the case in which both s and t are nuclear spin systems there exists in principle the same type of effect, provided the fluctuation of the lattice is large. In fact the coupled free induction in liquid hydrogen fluoride, which was observed by Solomon,[10] is an example of transient nuclear Overhauser effect. In this case we get no large polarization as in the case of electron-nucleus system ; however, the effect can be a convenient indicator of the resonance of the system to be saturated.

§ 5. B. *Example of saturational narrowing*

(1) Non-steady experiment

Corresponding to the case of rigid lattice we have a typical example in Herzog and Hahn's experiment[17] on $Na^{23}Cl^{35}O_3$. Under the continuous wave saturation of Na^{23}

resonance, they observed the non-steady behaviour of Cl^{35} nucleus, and established the fact that the transverse relaxation time T_2 of Cl^{35} increases on saturating Na^{23}. In the steady state language this corresponds to the narrowing of Cl^{35} resonance. Our results on Φ_{xy} are qualitatively similar to theirs ; however, they have assumed Φ_{zy} as constant, which actually changes according to the degree of saturation as we have discussed in detail, so that their analysis does not seem to be exact quantitatively.

(2) Steady state experiment

Saito[11] has observed a good example of saturational narrowing on various double resonance in thallous fluoride. Searls and Cotts[12] have also observed the same effect in

Table 3. Crystal structure of thallous fluoride (TlF).

Orthorhombic — 4-molecules in a unit cell					
a_0 5.180 Å	Tl	at	(000)	$(\frac{1}{2}\frac{1}{2}0)$	$(\frac{1}{2}0\frac{1}{2})$ $(0\frac{1}{2}\frac{1}{2})$
b_0 5.495 Å	F	at	$(\frac{1}{2}\frac{1}{2}\frac{1}{2})$	$(0\frac{1}{2}0)$	$(00\frac{1}{2})$ $(\frac{1}{2}00)$
c_0 6.080 Å					

$Na^{23}F^{19}$. We here propose to compare the result of our theory with Saito's data on $Tl^{205}F^{19}$. The information on the crystal structure is summarized in Table 3. and the nuclear properties are shown in Table 4, in which we see that actually there are two different isotopes, Tl^{205} and Tl^{203}.

Table 4. Nuclear data for TlF.

	F^{19}	Tl^{203}	Tl^{205}
(1) Resonance frequency (for 10^3 gauss)	4.0055 M. C.	2.433 M. C.	2.457 M. C.
(2) Natural abundance	100%	29.52%	70.48%
(3) Magnetic moment	2.6273 N. M.	1.5960 N. M.	1.6114 N. M.
(4) Spin	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$

(a) The case in which Tl^{205} is detected on saturating F^{19}.

In this case we have from (4·29)

$$\Phi_{\theta_2 y}=\sigma'^{\theta_1+\theta_2}_{\theta_2,00}+\sigma'^{F}_{\theta_2,00}\left\{1-e^{-\frac{\sigma''^{F}_{\theta_2,00}\Phi_{Fy}}{\Phi_{Fy}\Phi_{Fz}+\omega_1 F^2}}\right\}, \tag{5·1}$$

where the suffixes θ_1, θ_2 and F stand for Tl^{203}, Tl^{205} and F^{19}, respectively, and

$$\sigma^{\theta_1+\theta_2\,2}_{\theta_2,0}\equiv\sigma^{\theta_1\,2}_{\theta_2,00}+\sigma^{2}_{\theta_2,0}. \tag{5·2}$$

Under the assumption $\Phi_{\theta_1 z}=\Phi_{\theta_2 z}=\Phi_{\theta z}$, the Φ_{Fy} in (5·1) can be written

$$\Phi_{Fy}=\sigma'_{F,0}\left\{1-\frac{1}{2}e^{-\frac{\sigma_{F,0}\,''\,\phi_{F0}}{\phi_{F0}{}^2+\omega_1 F^2}}\right\}+\sigma'^{\theta_1+\theta_2}_{F,00}\left\{1-e^{-\frac{\sigma''^{\theta_1+\theta_2}_{F,00}\Phi_{\theta z}}{\Phi_{\theta z}{}^2+\omega_1 F^2}}\right\} \tag{5·3}$$

by using (4·24), where

$$\sigma_{F,00}^{\theta_1+\theta_2\,2} \equiv \sigma_{F,00}^{\theta_1\,2} + \sigma_{F,00}^{\theta_2\,2}\,. \tag{5·4}$$

If we suppose that the magnetic dipole interaction among fluorine nuclei makes negligible contribution to the lattice relaxation of fluorine, then we may put

$$\phi_{F0} = \Phi_{Fz} \tag{5·5}$$

as we did in paper I.

(b) The case in which F^{19} is detected on saturating T^{205}.

In this case we have from (4·29)

$$\Phi_{Fy} = \sigma'^{F+\theta_1}_{F,0} + \sigma'^{\theta_2}_{F,00}\left\{1 - e^{-\frac{\sigma''^{\theta_2}_{F,00}\Phi_{\theta_2 y}}{\Phi_{\theta_2 y}\Phi_{\theta_2 z}+\omega_{1\theta_2}{}^2}}\right\}, \tag{5·6}$$

where $\Phi_{\theta_2 y}$ is given by

Table 5. Calculated second moment (in the unit of gauss^2) for TlF.

Contribution from \ to	F^{19}	Tl^{205}
F^{19}	$\sigma_{F,0}{}^2 = 1.2222$	$\sigma^{F}_{\theta_2,00}{}^2 = 2.1189$
Tl^{203}	$\sigma^{\theta_1}_{F,00}{}^2 = 0.2308$	$\sigma^{\theta_1}_{\theta_2,00}{}^2 = 0.0592$
Tl^{205}	$\sigma^{\theta_2}_{F,00}{}^2 = 0.5619$	$\sigma_{\theta_2,0}{}^2 = 0.3258$
Total	2.0148	2.5038

Table 6. The values of σ', σ'' for TlF.

$\{\ ,\ \}$	σ' $\{\ ,\ \}$ gauss	σ'' $\{\ ,\ \}$ gauss
$\begin{Bmatrix}\theta_1+\theta_2\\ \theta_2,\ 0\end{Bmatrix}$	0.729	0.0528
$\begin{Bmatrix}F\\ \theta_2,\ 00\end{Bmatrix}$	1.1299	1.8752
$\begin{Bmatrix}\\ F,\ 0\end{Bmatrix}$	1.2990	0.9409
$\begin{Bmatrix}\theta_1+\theta_2\\ F,\ 00\end{Bmatrix}$	0.3689	2.1488
$\begin{Bmatrix}F+\theta_1\\ F,\ 0\end{Bmatrix}$	1.4164	1.0259
$\begin{Bmatrix}\theta_2\\ F,\ 00\end{Bmatrix}$	0.2516	2.2334
$\begin{Bmatrix}\\ \theta_2,\ 0\end{Bmatrix}$	0.6707	0.4857
$\begin{Bmatrix}\theta_1\\ \theta_2,\ 00\end{Bmatrix}$	0.2858	0.2070
$\begin{Bmatrix}F\\ \theta_2,\ 00\end{Bmatrix}$	0.9025	2.3478

Table 7. The value of Φ_z for TlF.

	cycles/sec	gauss
Φ_{Fz}	(2 msec→) 500	1.2484×10^{-1}
$\Phi_{\theta z}$	(50 msec→) 20	7.8524×10^{-3}

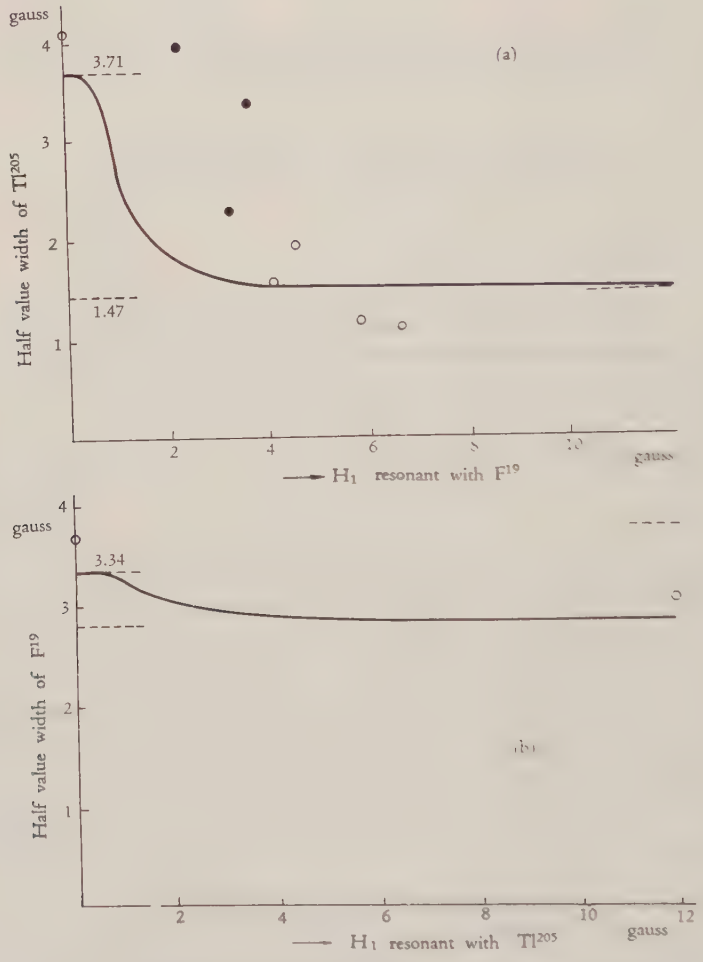


Fig. 3. Saturational narrowing of the nuclear resonance in TlF.
(a) The width of Tl^{205} on saturating F^{19}.
(b) The width of F^{19} on saturating Tl^{205}.

$$\Phi_{\theta_2 y}=\sigma'_{\theta_2,0}\left\{1-\frac{1}{2}e^{-\frac{\sigma''_{\theta_2 0}\phi_{\theta_2 0}}{\phi_{\theta_2 0}^2+\omega_{1\theta_2}^2}}\right\}+\sigma'^{\theta_1}_{\theta_2,00}\left\{1-e^{-\frac{\sigma''^{\theta_1}_{\theta_2,00}\phi_{\theta_1 0}}{\phi_{\theta_1 0}^2+\omega_{1\theta_2}^2}}\right\}$$

$$+\sigma'^{F}_{\theta_2,00}\left\{1-e^{-\frac{\sigma''^{F}_{\theta_2 00}\Phi_{Fz}}{\Phi_{Fz}^2+\omega_{1\theta_2}^2}}\right\}, \qquad (5\cdot7)$$

and we may put $\phi_{\theta_1 0}=\phi_{\theta_2 0}=\Phi_{\theta z}$ as a result of a similar consideration as before. The definitions of primed and doubly primed quantities have been given in the previous section; however, as all the above formulae are meant to be an interpolation, it is desirable to have a correct asymptotic values of width. For this purpose we slightly modify the definition of the quantities as follows.

$$\sigma'^{F}_{\theta_2,00}=1.175\{\sqrt{(\sigma^{\theta_1+\theta_2}_{\theta_2,0})^2+(\sigma^{F}_{\theta_2,00})^2}-\sigma^{\theta_1+\theta_2}_{\theta_2,0}\}, \tag{5·8}$$

$$\sigma'^{\theta_1+\theta_2}_{F,00}=1.175\{\sqrt{(\sigma_{F,0})^2+(\sigma^{\theta_1+\theta_2}_{F,00})^2}-\sigma_{F,0}\}, \tag{5·9}$$

$$\sigma'^{\theta_2}_{F,00}=1.175\{\sqrt{(\sigma^{F+\theta_1}_{F,0})^2+(\sigma^{\theta_2}_{F,00})^2}-\sigma^{F+\theta_1}_{F,0}\}, \tag{5·10}$$

$$\sigma'^{F}_{\theta_2,00}=1.175\{\sqrt{(\sigma_{\theta_2,0})^2+(\sigma^{\theta_1}_{\theta_2,00})^2+(\sigma^{F}_{\theta_2,00})^2}-(\sigma_{\theta_2,0}+\sigma^{\theta_1}_{\theta_2,00})\}. \tag{5·11}$$

Other quantities are defined by the standard formulae. The calculated values of necessary second moments are given in Table 5, and those of σ' and σ'' are given in Table 6. By using the values of observed lattice relaxations, which is given in Table 7, the theoretical results are compared with the observation by Saito in Fig. 3. It should be emphasized that no adjustable parameters were employed. We see that the general tendency is satisfactorily reproduced; however, quantitatively the agreement is not quite exact. On the theoretical side the cause of this discrepancy may be traced back to i) the unproven legitimacy of the interpolation formulae employed, and ii) the use of the phenomenological relaxation Φ instead of ϕ.

Experimentally there appears a skewness and a collapse of the absorption of the traced nucleus in the early stage of saturation, which is not simply derived from our theory. Bloembergen[18] has recently argued that this phenomenon can be explained by the higher order transitions using the results of simple two spin system. If this is the case, we have neglected such effect at the stage in which we have factorized the spin density matrix into two parts and have taken an average over one kind in Hartree's sense, i. e. assuming the density as diagonal. The neglected off-diagonal part is responsible for the higher order transitions, therefore it is only natural that we have missed the explanation. It is desirable to check the Bloembergen's explanation quantitatively for a real example consisting of many spins, which may be amenable to the moment method.

In the expression (4·12) of the longitudinal relaxation Ψ_{tz} the term $\Psi'^{11\bar{1}}*$ corresponds to that which was discussed by Kaplan.[19),20)] In the case of rigid lattice, however, $\Psi'^{11\bar{1}}*$ is explicitly given by

$$\Psi'^{11\bar{1}}*=\sigma^2_{t11}\int_0^t d\tau\, e^{i(\omega_s-\omega_t)\tau} f_1^{s}*(\tau), \tag{5·12}$$

so that we should use the relaxation function $f_1^{s}*(\tau)$ in the rotating system. Kaplan himself has simply assumed

$$f_{\bar{1}}^{s}*(\tau)=e^{-\frac{t}{T_{rf}{}^{el}}}\ ;\ (T_{rf}{}^{el})^{-1}=\tfrac{1}{2}\mu_B{}^2 H_1{}^2 T_2{}^{el}/\hbar^2, \tag{5·13}$$

but there seems to be no justification, for we obtain a different function based on the

theory as will be shown (A·10) and (A·8). In addition $\Psi'^{(1)*}$ cannot be a dominant part, as Kaplan presumed, when there is an appreciable contribution of $\Psi_t'^{(1)}$ which is inherent to the t-system, and Ψ_{tz} is little affected by saturating the s-system.

The method of double resonance has several merits and their applications in the case of rigid lattice.

(1) When we are in search of a weak resonance of the s-system, which is difficult to observe owing to the weakness of absorption, etc., the narrowing of the t-system resonance can be used as an indicator for the s-resonance as was shown by Herzog and Hahn.[1)] The point is that we need no appreciable population difference of spin levels of the s-system. A mere communication between s-levels which is enhanced by the resonance is enough to modify the transverse relaxation time, or the width of the t-system.

(2) When there are several different origins of the local field being felt by a nucleus, the method of double resonance may separate out the effect of just one kind of the origins quantitatively. Therefore the method will doubtlessly be helpful in the detailed analysis of the local field.

Finally we might mention the work of Bloom and Shoolery[21)] which has much to do with the physics presented in this paper. They observed a collection of line spectra, not a single hump as we postulated in this paper, but the behaviour of each line upon saturation is quite suggestive. It should be emphasized that in their observation, too, the central components are enhanced at the expense of the wing components in the t-system resonance on saturating the s-system. It is desirable to make a detailed analysis of simple systems[22),23)] to look into further details of the mechanism of relaxation, which will, however, require a more detailed theoretical treatment than is given here.

§ 6. Conclusion

Double resonance is one of the most ingeneous methods to analyse and also artificially to modify the local situation surrounding a magnetic moment; however, the observed characteristics varies delicately from case to case and we cannot expect a universal result, as once there was a conceptual objection to the Overhauser polarization effect. In this paper it has been clarified under what criteria we observe the Overhauser effect on the one hand, and the saturational narrowing on the other. We found that the situation solely depends on the relative magnitude of the strength of the static local field σ_{α} due to the interaction of the systems and the frequency ω_{α} with which the interaction is modulated by the environments. A) When the fluctuation of the environment dominates, i. e. $\sigma_{\alpha} \ll \omega_{\alpha}$, it becomes an energy reservoir, so that it absorbs the excess energy instantaneously. Suppose the s-system is excited by the radiation, the t-system is polarized through the interaction in a definite sence, but before the s-system re-emits the radiation the environment takes away the excess energy, thus leaving the t-system polarized in a definite sense. In other words, the interaction plays the role of an energy pump. This is the case where the Overhauser polarization effect is observed. B) When, on the other hand, the relative precession due to the interaction dominates, i. e. $\sigma_{\alpha} \gg \omega_{\alpha}$, the interaction is considered

just as a rigid mechanism connecting the states of the two systems in an unambiguous way. Even if the *s*-system is excited by the radiation, it eventually re-emits the same amount of energy quantum in the form of radiation. Side by side the polarization of the *t*-system is also flipped in an analogous way. This means that upon saturating the *s*-system the *t*-system is also saturated by a forced modulation. This situation is quite different from the previous case. Also at the location of the *t*-system the local field variety is diminished owing to the rapid nutation of the *s*-system, thus resulting in the saturational narrowing.

Acknowledgements

The author would like to thank Prof. T. Kanda, Drs Y. Masuda and Y. Saito who have kindly supplied him with interesting experimental data prior to publication. He would also thank Prof. J. Itoh, Prof. R. Kubo, and the members of the Solid State Physics Group in Kyoto for their special interest in this work and for their valuable comments and their stimulating discussion.

Appendix A

The case of spin $I=1/2$.

When the fluctuation of the lattice is dominant the equation for the *t*-system becomes automatically linear in spin components for the special case of spin $I=\frac{1}{2}$. In particular the *z*-components obey the following equation.

$$\begin{aligned}\frac{d}{dt}\langle I_{tz}\rangle = &-2\big[(\Phi'^{1\bar{1}}+\tfrac{1}{4}\Phi'^{01})\Big\{(\tfrac{1}{2}+\langle I_{tz}\rangle)-e^{\frac{\hbar\omega_t}{kT}}(\tfrac{1}{2}-\langle I_{tz}\rangle)\Big\}\\ &+\Phi'^{1\bar{1}}\Big\{(\tfrac{1}{2}-\langle I_{sz}\rangle)(\tfrac{1}{2}+\langle I_{tz}\rangle)-e^{\frac{\hbar(\omega_t-\omega_s)}{kT}}(\tfrac{1}{2}+\langle I_{sz}\rangle)(\tfrac{1}{2}-\langle I_{tz}\rangle)\Big\}\\ &+\Phi'^{11}\Big\{(\tfrac{1}{2}+\langle I_{sz}\rangle)(\tfrac{1}{2}+\langle I_{tz}\rangle)-e^{\frac{\hbar(\omega_t+\omega_s)}{kT}}(\tfrac{1}{2}-\langle I_{sz}\rangle)(\tfrac{1}{2}-\langle I_{tz}\rangle)\Big\}\\ \equiv &\Phi_{tz}[\langle I_{tz}\rangle_0-\langle I_{tz}\rangle],\end{aligned}\tag{A·1}$$

where

$$\begin{aligned}\Phi_{tz}=2\Big[&(\Phi'^{1}+\tfrac{1}{4}\Phi'^{01})(e^{\frac{\hbar\omega_t}{kT}}+1)\\ &+\tfrac{1}{2}\Big\{\Phi'^{1\bar{1}}(e^{\frac{\hbar(\omega_t-\omega_s)}{kT}}+1)+\Phi'^{11}(e^{\frac{\hbar(\omega_t+\omega_s)}{kT}}+1)\Big\}\\ &+\langle I_{sz}\rangle\Big\{\Phi'^{1\bar{1}}(e^{\frac{\hbar(\omega_t-\omega_s)}{kT}}-1)-\Phi'^{11}(e^{\frac{\hbar(\omega_t+\omega_s)}{kT}}-1)\Big\}\Big],\end{aligned}\tag{A·2}$$

and

$$\begin{aligned}\langle I_{tz}\rangle_0\Phi_{tz}=\Big[&(\Phi'^{1}+\tfrac{1}{4}\Phi'^{01})(e^{\frac{\hbar\omega_t}{kT}}-1)+\tfrac{1}{2}\Big\{\Phi'^{1\bar{1}}(e^{\frac{\hbar(\omega_t-\omega_s)}{kT}}-1)+\Phi'^{11}(e^{\frac{\hbar(\omega_t+\omega_s)}{kT}}-1)\Big\}\\ &+\langle I_{sz}\rangle\Big\{\Phi'^{1\bar{1}}(e^{\frac{\hbar(\omega_t-\omega_s)}{kT}}+1)-\Phi'^{11}(e^{\frac{\hbar(\omega_t+\omega_s)}{kT}}+1)\Big\}\Big].\end{aligned}\tag{A·3}$$

Appendix B

The solution of the Bloch equation in the rotating frame.

At exact resonance the Bloch equation in the rotating frame is given by

$$\begin{cases} \dot{u}+\Phi_x u=\Phi_x u_0, & (A\cdot 4)\\ \dot{v}+\Phi_y v=-\omega_1 M_z, & (A\cdot 5)\\ \dot{M}_z+\Phi_z M_z=\omega_1 v+\Phi_z M_0. & (A\cdot 6) \end{cases}$$

Eliminating v from (A·5) and (A·6) we have

$$\ddot{M}_z+(\Phi_z+\Phi_y)\dot{M}_z+(\Phi_z\Phi_y+\omega_1^2)M_z=(\Phi_z\Phi_y)M_0. \qquad (A\cdot 7)$$

Putting $M_z\propto e^{-\Phi t}$, we can solve the relaxation constant Φ as

$$\begin{pmatrix}\Phi_2^-\\ \Phi_1^+\end{pmatrix}=\frac{1}{2}\{(\Phi_y+\Phi_z)\pm\sqrt{(\Phi_y-\Phi_z)^2-4\omega_1^2}\}, \qquad (A\cdot 8)$$

so that the solution becomes oscillatory when $\omega_1 > |(\Phi_y-\Phi_z)/2|$. (A·8) corresponds to (4·26) in the text. The solution of the inhomogeneous equation is given by

$$M_z(t)=c_1 e^{-\Phi_2^- t}+c_2 e^{-\Phi_1^+ t}+M_0\left(\frac{\Phi_y\Phi_z}{\Phi_y\Phi_z+\omega_1^2}\right) \qquad (A\cdot 9)$$

$$v(t)=\frac{1}{\omega_1}\{(\Phi_z-\Phi_2^-)c_1 e^{-\Phi_2^- t}+(\Phi_z-\Phi_1^+)c_2 e^{-\Phi_1^+ t}\}-M_0\left(\frac{\omega_1\Phi_z}{\Phi_y\Phi_z+\omega_1^2}\right) \qquad (A\cdot 10)$$

$$u(t)=ce^{-\Phi_2 t}+u_s. \qquad (A\cdot 11)$$

Table 8. Non-steady behaviour of the polarization in the rotating frame upon saturation.

	Initial condition	The behaviour at later time	
I	$\Delta u=\Delta u_0$	$\rightarrow \Delta u=\Delta u_0 e^{-\Phi_x t}=\Delta u_0 f_x^*(t)$	(A·12)
	$\Delta v=0$	$\rightarrow \Delta v=0$	(A·13)
	$\Delta M_z=0$	$\rightarrow \Delta M_z=0$	(A·14)
II	$\Delta u=0$	$\rightarrow \Delta u=0$	(A·15)
	$\Delta v=\Delta v_0$	$\rightarrow \Delta v=\Delta v_0\dfrac{(\Phi_2^- -\Phi_z)e^{-\Phi_2^- t}+(\Phi_z-\Phi_1^+)e^{-\Phi_1^+ t}}{\Phi_2^- -\Phi_1^+}=\Delta v_0 f_y^*(t)$	(A·16)
	$\Delta M_z=0$	$\rightarrow \Delta M_z=\Delta v_0\dfrac{\omega_1(-e^{-\Phi_2^- t}+e^{-\Phi_1^+ t})}{\Phi_2^- -\Phi_1^+}$	(A·17)
III	$\Delta u=0$	$\rightarrow \Delta u=0$	(A·18)
	$\Delta v=0$	$\rightarrow \Delta v=\Delta M_{z0}\dfrac{(\Phi_z-\Phi_2^-)(\Phi_z-\Phi_1^+)(e^{-\Phi_2^- t}-e^{-\Phi_1^+ t})}{\omega_1(\Phi_2^- -\Phi_1^+)}$	(A·19)
	$\Delta M_z=\Delta M_{z0}$	$\rightarrow \Delta M_0-\Delta M_{z0}\dfrac{(\Phi_z-\Phi_1^+)e^{-\Phi_2^- t}-(\Phi_z-\Phi_2^-)e^{-\Phi_1^+ t}}{\Phi_2^- -\Phi_1^+}=\Delta M_{z0} f_z^*(t)$	(A·20)

If we denote the time dependent part by using J, we have the results shown in Table 8, corresponding to several different initial conditions.

The Fourier transforms of $f_x^*(t)$, $f_y^*(t)$, and $f_z^*(t)$ are given by

$$\int_0^\infty f_x^*(t)\, e^{i\Omega t} dt = \frac{\Phi_x}{\Omega^2 + \Phi_x^2} \tag{A·21}$$

$$\int_0^\infty f_y^*(t)\, e^{i\Omega t} dt = \frac{\{\Omega^2 \Phi_y + \Phi_z(\Phi_z\Phi_y + \omega_1^2)\} + i\Omega\{\Omega^2 + (\Phi_z^2 - \omega_1^2)\}}{\Omega^4 + \Omega^2(\Phi_z^2 + \Phi_y^2 - 2\omega_1^2) + (\Phi_y\Phi_z + \omega_1^2)^2} \tag{A·22}$$

$$\int_0^\infty f_z^*(t)\, e^{i\Omega t} dt = \frac{\{\Omega^2 \Phi_z + \Phi_y(\Phi_y\Phi_z + \omega_1^2)\} + i\Omega\{\Omega^2 + (\Phi_y^2 - \omega_1^2)\}}{\Omega^4 + \Omega^2(\Phi_z^2 + \Phi_y^2 - 2\omega_1^2) + (\Phi_y\Phi_z + \omega_1^2)^2}. \tag{A·23}$$

If we take the limit of $\Omega \to 0$ in the real part of the last equation, we immediately obtain (4·27) in the text.

References

1) A. W. Overhauser, Phys. Rev. **91** (1953), 476; **92** (1953), 411.
2) F. Bloch, Phys. Rev. **93** (1954), 944 (A).
3) J. Korringa, Phys. Rev. **94** (1954), 1388.
4) C. Kittel, Phys. Rev. **95** (1954), 589.
5) A. Abragam, Phys. Rev. **98** (1955), 1729.
6) M. J. Klein, Phys. Rev. **98** (1955), 1736.
7) J. I. Kaplan, Phys. Rev. **99** (1955), 1322.
8) T. R. Carver and C. P. Slichter, Phys. Rev. **102** (1956), 975.
9) W. A. Barker and A. Mencher, Phys. Rev. **102** (1956), 1023.
10) I. Solomon, Phys. Rev. **99** (1955), 559.
11) Y. Saito, J. Phys. Soc. Japan **13** (1958), 72.
12) L. R. Searles and R. M. Cotts, Bull. Am. Phys. Soc. p. 282; Phys. Rev. **111** (1958), 853.
13) B. Herzog and E. L. Hahn, Phys. Rev. **103** (1956), 148.
14) K. Tomita, Prog. Theor. Phys. **19** (1958), 541.
15) W. D. Knight, Phys. Rev. **76** (1949), 1259.
16) C. P. Slichter, Phys. Rev. **99** (1955), 1822; ibid **103** (1956), 1905.
17) H. Hasegawa, to be published.
18) N. Bloembergen, J. Phys. Soc. Japan **13** (1958), 660.
19) A. Honig, Phys. Rev. **96** (1954), 234.
20) J. I. Kaplan, Phys. Rev. **96** (1954), 235.
21) A. L. Bloom and J. N. Shoolery, Phys. Rev. **97** (1957), 1261.
22) J. Itoh and collaborators, private communication.
23) N. Bloembergen, Phys. Rev. **104** (1956), 1542.

Note added in proof. Bloch has put forth a theory of line narrowing in a recent paper (Phys. Rev. **111** (1958), 841), which the writer has had a chance to look at but has had little time to study carefully before he sends in his final proof. His treatment seems to correspond to the situation B, i.e. the case of *Saturational narrowing*, according to the classification in the present paper. The concept of saturational narrowing corresponds to Bloch's recognition of the growth of sidebands at the expense of central component. It should be remarked, however, that in general not only the relaxation times of the traced system but also those of the saturated system are changed upon saturation. This latter effect has already been treated as the situation B in paper I by the present writer, and its existence makes difference when we try to describe the intermmediate region corresponding to a partial narrowing.

Letters to the Edior

The opinions expressed in these columns do not necessarily reflect those of the Board of Editors. Communications should be submitted in duplicate and should be held to within 100 lines (pica type) on standard size letter paper (approx. 21×30cm.), so that each letter will be arranged into two pages when printed. Do not forget to count in enough space for formulas, figures or tables.

On the Absorptive Range in the Pion-Nucleon Collision at High Energies

Kin-ichi Ishida

Faculty of Liberal Arts and Science, Yamagata University, Yamagata

August 8, 1958

According to the recent experiments on the pion-nucleon collision at high energies with an additional pion production,[1] (1) the angular distribution of the recoil nucleon has a peak in the backward direction in the center of mass system, (2) the angular distribution relative to the recoil nucleon of the more energetic pion has a strong peak in the forward direction and that of the slower is nearly isotropic. It has been shown by Kovacs[2] and extensively by Ito et al.[3] that these evidences seem to suggest the knock-out process in which the incident pion interacts strongly with the pion field surrounding the nucleon core, and knocks out the pions, rather than to suggest the thermal process in which the incident pion is absorbed directly by the core. This may also be supported by the fact that the elastic scattering is consistent with the diffraction scattering by an absorptive disk with a radius of about $(1.0\sim1.2)\times10^{-13}$ cm,[1] because this large radius does not seem to be due to the direct absorption by the core. Here, we shall then investigate the absorptive range from the standpoint of the $\pi-\pi$ interaction model.

In the low energy region, the absorptive phenomena may be described by the imaginary potential $iW(r)$ as the function of the relative coordinate of the interacting particles. In the relativistic region, however, such a simple picture may not hold, so that it is necessary to introduce a new physical measure to specify the absorptive range. For this purpose, it is convenient to define the root mean square $(\bar{d}^2)^{1/2}$ of the impact parameter of the incident pion relative to the nucleon, in such a way as

$$p^2\bar{d}^2=\bar{l}^2=\sum_l l(l+1)\sigma_l \Big/ \sum_l \sigma_l,^{4} \qquad (1)$$

where σ_l is the cross section of the partial wave with the orbital angular momentum l, and $\boldsymbol{P}$ the relative momentum. The right-hand side of eq. (1) is simplified on account of the angular momentum conservation,[5] and is expressed as

$$\bar{d}^2=\frac{\{L'^2\langle \boldsymbol{P}'|T^{\dagger}\delta(E-H_0)T|\boldsymbol{P}\rangle\}_{\boldsymbol{P}'\to\boldsymbol{P}}}{p^2\langle \boldsymbol{P}|T^{\dagger}\delta(E-H_0)T|\boldsymbol{P}\rangle}, \qquad (2)$$

where T is the reaction matrix, $|\boldsymbol{P}\rangle$ and $|\boldsymbol{P}'\rangle$ are the initial and final states with the relative momentum $\boldsymbol{P}$ and $\boldsymbol{P}'$ $(p=p')$, L'^2 is the square of the angular momentum operator acting on $\boldsymbol{P}'$, E is the total energy of the system, and H_0 the free Hamiltonian. $\bar{d}^2$ thus introduced is the simplest quantity, because it is unnecessary to calculate each σ_l.

We then calculate the $\bar{d^2}$ under the following assumptions. The pion number surrounding the core is only one, which may be supported by the fixed source theory, at least, in the outer region in which the knock-out process may be dominant; the pion-core interaction Hamiltonian is

$$H=i\,g/\mu\cdot\tau_i\sigma\cdot\mathrm{grad}\,\phi_i(\theta), \qquad (3)$$

when the nucleon is assumed to be fixed in the laboratory system; the local $\pi-\pi$ interaction Hamiltonian is simply

$$H_{\pi\pi}=\int\lambda\phi^4(x)\,d^3x\,; \qquad (4)$$

finally, the final state interactions between the outgoing particles can be neglected, which seems to be natural for the knock-out process. The calculation is straightforward. The result is simplified in the high energy limit as follows,

$$\bar{d^2}=\pi\hbar^2/2m_\pi cp, \quad \text{for } \hbar/p\ll\hbar/m_\pi c, \qquad (5)$$

where p is the pion momentum in the laboratory system, and m_π is the pion mass. The numerical values are listed in Table 1. As the uniform absorptive disk with the radius R gives $\bar{d^2}=R^2/2$, the calculated $\bar{d^2}$ is smaller compared with the experimentally inferred one.[1] (If one assumes the scalar type interaction $H=f\phi$, the $(\bar{d^2})^{1/2}$ is calculated as $\hbar/\sqrt{3}\,m_\pi c\fallingdotseq 0.82\times10^{-13}$ cm, which is independent of the p.)

Table 1. The root mean square of the impact parameter, $(\bar{d^2})^{1/2}$ and the radius of the uniform absorptive disk, R $(\bar{d^2}=R/2)$.

Energy of incident pion	1.0 Bev	1.4 Bev	2.0 Bev
$(\bar{d^2})^{1/2}\times10^{13}$ cm	0.67	0.56	0.48
$R\times10^{13}$ cm	0.94	0.78	0.67

The $\pi-\pi$ interaction model, however, may not be inconsistent with the experiments, if one takes into account the nucleon recoil correction and the spread-out $\pi-\pi$ interaction. Since the local $\lambda\phi^4$ interaction which acts only between the isotopic spin even states may not be consistent with the experimental charge ratios of the outgoing pions, the $\lambda\phi^4$ thus may play little role except to cancell the divergences in the renormalizable field theory, and the correct "Hamiltonian" may rather be expressed as the spread-out interactions including the space-time derivatives.[6] Details of this note will be reported in near future. The author would like to thank Prof. Nakabayasi for his encouragement and many valuable discussions. He also thanks the members of Tohoku University for helpful discussions.

1) W. D. Walker and J. Crussard, Phys. Rev. **98** (1955), 1416. L. M. Eisberg, W. B. Fowler, R. M. Lea, W. D. Shephard, R. S. Shutt, A. M. Thorndike, and W. L. Whittemore, Phys. Rev. **97** (1955), 797. G. Maenchen, W. B. Fowler, W. M. Powell, and R. W. Wright, Phys. Rev. **108** (1957), 850. W. D. Walker, Phys. Rev. **108** (1957), 872.
2) J. S. Kovacs, Phys. Rev. **93** (1954), 252.
3) Ito, Yamazaki, and Mori, Soryushiron Kenkyu (in Japanese) **10** (1956), 259.
4) The similar quantity has been discussed independently by Ito et al.; Soryushiron Kenkyu (in Japanese) **17** (1958), 19.
5) We simplified the discussions averaging over the nucleon spin directions.
6) Kawarabayashi and Igi, Soryushiron Kenkyu (in Japanese) **16** (1957), 381.

An Illustrative Example of the Nuclear Collective Motion

Fumiaki Iwamoto

Department of Physics, Rikkyo University, Tokyo

August 20, 1958

Recentey Elliott and others[1] have shown that states of the rotational property emerge with a particular coupling among shell modell configurations. In order to see the underlying mechanism in a clear-cut way it may be of some interest to study the behavior of another simple dynamical system with quadrupole interactions among particles.

Instead of many fermions in many levels let us suppose a system of bosons in a two-dimensional well and consider just two levels, i. e., the lower s-level with eigenfunction $R_0(r)$ and the higher d-level $R_2(r)\,e^{\pm 2i\theta}$. Denoting the creation and annihilation operators of these states by b_m^*, b_m $(m=0,\ \pm 2)$, the Hamiltonian of the system is assumed to be

$$\begin{aligned} H=&T(b_2^*b_2+b_{-2}^*b_{-2})\\ &-V(b_2^*b_0^*b_2b_0+b_{-2}^*b_0^*b_{-2}b_0\\ &+b_2^*b_{-2}^*b_0b_0+b_0^*b_0^*b_2b_{-2}), \end{aligned} \tag{1}$$

where T is the d-level energy measured from the s-level and V the matrix element with angular momentum transfer by 2 units. This attractive quadrupole interaction is the essential character of the model which, although quite unrealistic, exhibits main features of nuclear collective motions.

Since the total number N is a constant of motion, the Hamiltonian is expressed by two sets of creation annihilation operators;

$$\begin{aligned} &\beta_m^*=b_m^*b_0(b_0^*b_0)^{-1/2},\\ &\beta_m=(b_0^*b_0)^{-1/2}b_0^*b_m, \quad (m=\pm 2). \end{aligned} \tag{2}$$

Or, changing to canonical variables

$$\begin{aligned} &p_m=(i/\sqrt{2})\ (\beta_m^*-\beta_{-m}),\\ &q_m=(1/\sqrt{2})\ (\beta_m+\beta_{-m}^*), \end{aligned} \tag{3}$$

$$\begin{aligned} &[p_m,\ q_{m'}]=-i\delta_{mm'},\\ &[p_m,\ p_{m'}]=[q_m,\ q_{m'}]=0, \quad (m,\ m'=\pm 2), \end{aligned}$$

we can write

$$\begin{aligned} H=&Tp_2p_{-2}+(T-2NV)q_2q_{-2}\\ &+2V(p_2p_{-2}+q_2q_{-2})q_2q_{-2}\\ &-(T-NV), \end{aligned} \tag{4}$$

where terms of $0(1/N)$ are neglected.

As is often done in other problems, we shall find the approximate solution as small oscillations about the equilibrium point. There are two cases of different characters according as $T-2NV\gtrless 0$.

As long as $T-2NV>0$, the minimum of the Hamiltonian stays at $p_m=q_m=0$. Then, we neglect the fourth order terms of p, q,[2] and the Hamiltonian (4) becomes a two-dimensional harmonic oscillator with the energy spectrum:

$$\begin{aligned} &E_{n_2,\,n_{-2}}=(n_2+n_{-2}+1)\omega-(T-NV),\\ &n_2,\ n_{-2}=0,\ 1,\ 2,\ \cdots, \end{aligned} \tag{5}$$

where the eigenfrequency ω is $1/T(T-2NV)$, which is smaller than the inititial separation T between s- and d-levels. The ground state wave function is

$$\Psi_0=\exp\left(-\frac{\omega-T}{T}q_2q_{-2}\right)\cdot\Phi_0, \tag{6}$$

where Φ_0 is the free ground state and the exponential factor represents the correlations induced by the interaction V. This is the weak coupling case corresponding to the

vibrational levels of medium weight nuclei.

In the strong coupling case, $T-2NV<0$, the minimum forms a circle on the q-plane, i. e., $p_m=0$ and $q_2q_{-2}=\beta_0^2$, where $\beta_0=\sqrt{(2NV-T)/4V}$. As was done by Wentzel[3] in his strong coupling meson theory, it is convenient to transform q's to polar coordinate by

$$q_2=\beta e^{-2i\phi}, \quad q_{-2}=\beta e^{2i\phi}. \tag{7}$$

Then p_2p_{-2} is the two-dimensional Laplace operator and p_ϕ is, in the original variables, $2(b_2^*b_2-b_{-2}^*b_{-2})$ which is the total angular momentum operator with eigenvalues $M=0, \pm2, \pm4, \cdots$. For the small displacement around the minimum at $\beta=\beta_0$ the Hamiltonian is approximated as

$$\begin{aligned}H=&\left(\frac{T}{8}+\frac{NV}{4}\right)\\&\times\left\{\frac{1}{\sqrt{\beta}}p_\beta^2\sqrt{\beta}+\frac{1}{4\beta_0^2}(p_\phi^2-1)\right\}\\&+8V\beta_0^2(\beta-\beta_0)^2-2V\beta_0^4\end{aligned} \tag{8}$$

Its eigenvalues and eigenfunctions are

$$\begin{aligned}E_{n\beta M}=&\left(n_\beta+\frac{1}{2}\right)\omega_\beta\\&+\frac{1}{2\vartheta}(M^2-1)-2V\beta_0^4,\\\Psi_{n\beta M}=&\frac{1}{\sqrt{\beta}}H_{n\beta}(\sqrt[4]{4V\vartheta}(\beta-\beta_0))\\&\times e^{-\sqrt{V}\vartheta(\beta-\beta_0)^2}\cdot e^{iM\phi},\end{aligned} \tag{9}$$

where $\omega_\beta=2\beta_0\sqrt{V(T+2NV)}$ and $\vartheta=16\beta_0^2/(T+2NV)$. These spectra correspond to the rotation-vibration levels in heavy nuclei.

It is to be noted that in the configuration space our q_2, q_{-2} may be written approximately as

$$\left.\begin{matrix}q_2\\q_{-2}\end{matrix}\right\}\rightarrow\sum_i f(r_i)\,e^{\mp2i\theta_i}, \tag{10}$$

which are the collective coordinates in the usual theory[4] except that $f(r)$ is not simple r^2, but depends on the radial functions $R_0(r)$, $R_2(r)$. With these correspondence it is easy to show that β represents the deformation and ϕ the orientation of our system.

1) J. P. Elliott, Proc. Roy. Soc. **A245** (1958), 128. S. A. Moszkowski, Phys. Rev. **110** (1958), 403.
2) This corresponds to the approximation first used by Bogolyubov in the theory of the phonon state of liquid helium: N. Bogolyubov, J. Phys. U. S. S. R. **11** (1947), 23.
3) G. Wentzel, Helv. Phys. Acta **13** (1940), 269.
4) S. Tomonaga, Prog. Theor. Phys. **13** (1955), 467.

Molecular Processes induced by μ^- Mesons in Hydrogen Bubble Chamber. I†

——Transfer Process of μ^- Meson from Proton to Deuteron——

M. Shimizu, Y. Mizuno and T. Izuyama

Department of Physics, University of Tokyo, Tokyo

August 28, 1958

We here report our study on the process of transfer of a μ^- meson from a proton to a deuteron, when a μ^- mesic proton

† This is one of our reports which were published in the *Progress Report of Research Group for the Study of Molecular Structure in Japan*, No. 7 (December, 1957).

atom collides with a deuteron with very low incident energies.[1] Estimation of the reaction time of this process has been attempted by many people[2] and they have found the values which are spread over a rather wide range. We carried out a calculation which stands on a sounder basis than the previous ones and is expected to give a better result.

The total Hamiltonian after separating the motion of center of gravity is given by

$$H=-\frac{\hbar^2}{2m}\Delta_r-\frac{e^2}{r_p}-\frac{e^2}{r_d}+\frac{e^2}{R}-\frac{\hbar^2}{2M}\Delta_R \qquad (1)$$

where $\boldsymbol{R}$ is the relative vector from p to d, $\boldsymbol{r}$ is that from the center of gravity of p and d to μ^-, r_p and r_d are the distances between μ^- and p and d, respectively, and, further, M and m are respectively $(1/M_p+1/M_d)^{-1}$ and $(1/m_\mu+1/(M_p+M_d))^{-1}$.

Here, we assume the required wave function to have the form

$$\Psi=g(\boldsymbol{r},\boldsymbol{R})G(\boldsymbol{R})+u(\boldsymbol{r},\boldsymbol{R})U(\boldsymbol{R}). \qquad (2)$$

In the above, g and u are respectively $(1s\sigma)$ and $(2p\sigma)$ orbitals satisfying the equation

$$\left(-\frac{\hbar^2}{2m}\Delta_r-\frac{e^2}{r_p}-\frac{e^2}{r_d}+\frac{e^2}{R}\right)\varphi(\boldsymbol{r},\boldsymbol{R})=\varepsilon(R)\varphi(\boldsymbol{r},\boldsymbol{R}), \qquad (3)$$

in which $\boldsymbol{R}$ appears only as a parameter. Their asymptotic forms as $R\to\infty$ are as follows:

$$g=\frac{1}{\sqrt{2}}(h(r_p)+h(r_d))$$

and

$$u=\frac{1}{\sqrt{2}}(h(r_p)-h(r_d)), \qquad (4)$$

$h(r)$ being the $1s$ orbital of the μ mesic hydrogen atom with the reduced mass of m. Then the variation principle leads to the simultaneous equations for G and U,

$$\Delta_R G+\frac{2M}{\hbar^2}(W-\varepsilon_g)G+\left(\int g^*\Delta_R g\,d\boldsymbol{r}\right)G=-\left(\int g^*\operatorname{grad}_R u\,d\boldsymbol{r}\right)\cdot\operatorname{grad}_R U-\left(\int g^*\Delta_R u\,d\boldsymbol{r}\right)U \qquad (5)$$

$$\Delta_R U+\frac{2M}{\hbar^2}(W-\varepsilon_u)U+\left(\int u^*\Delta_R u\,d\boldsymbol{r}\right)U=-\left(\int u^*\operatorname{grad}_R g\,d\boldsymbol{r}\right)\cdot\operatorname{grad}_R G-\left(\int u^*\Delta_R g\,d\boldsymbol{r}\right)G,$$

where W is the total energy associated with the Hamiltonian of (1).

Accordingly, our task is to find the solution of eq. (5) satisfying the boundary condition as $R\to\infty$

$$\Phi(\boldsymbol{r},\boldsymbol{R})\to e^{i\boldsymbol{k}\cdot\boldsymbol{R}}h(r_p)+F\frac{e^{ik'R}}{R}h(r_p)+F'\frac{e^{ik'R}}{R}h(r_d). \qquad (6)$$

Now eq. (5) is invariant with respect to any rotation of the $\boldsymbol{R}$ space. Then we can solve eq. (5) by the method of partial waves. Further, since we consider only the case of very small values of $|\boldsymbol{k}|$, we may take up only the s wave. For a given value of $\boldsymbol{k}$, the values of k', F, F' and W can be determined by the requirement that the condition (6) should be compatible with eq. (5). As a result of this requirement we get the relation

$$\frac{\hbar^2k'^2}{2M}-\frac{\hbar^2k^2}{2M}=\frac{1}{2M}\int h(r_d)\Delta_R h(r_d)\,d\boldsymbol{r}-\frac{1}{2M}\int h(r_p)\Delta_R h(r_p)\,d\boldsymbol{r}=159\text{ e. v.}$$

Table I. Reaction Time in Normal Hydrogen Bubble Chamber* (T)

Author	Skyrme	Jackson	Ours	Cohen et al.[4]	Hayashi et al.
T (in 10^{-6} sec)	0.01	0.02	0.39	0.62?	3

* The normal hydrogen bubble chamber contains 5.9×10^{18} deuterons per cm^3.

Then we find the energy Q released by the present reaction is

$$Q=\frac{1}{2}\left(\frac{1}{M_d+m_\mu}+\frac{1}{M_p}\right)\hbar^2k'^2$$
$$-\frac{1}{2}\left(\frac{1}{M_p+m_\mu}+\frac{1}{M_n}\right)\hbar^2k^2$$
$$=\left[148-0.0748\frac{\hbar^2k^2}{2M_p}\right]\text{e. V.}$$

Especially, for the collision with cold velocity under consideration, this may be safely replaced by

$$Q=148\text{ eV},$$

while the corresponding correct value or experimental value is 135 eV and is of course completely independent of k. Then this result seems to suggest that the approximation (2) is a fairly satisfactory one even for μ^- mesons, if the nuclear motion is sufficiently slow, as well as for electrons.

For ε_g and ε_u of eq. (5), we have adopted the ones calculated by Teller.[3] On the other hand, the non-adiabatic terms or the kinetic correction terms have been evaluated for large R by using LCAO approximation and for small R by using UAO approximation for the orbitals g and u. Then two independent solutions of eq. (5) have been found by integrating it numerically up to $R=6$ (in μ mesic unit) and the required solution composed of these two solutions has been determined by the requirement that it must smoothly join with the approximate analytic solution for $R>6$ which has the form of an s wave of eq. (6).

In Table I, the reaction time found by this calculation is given together with those by other authors.

1) L. W. Alvalez, et al., Phys. Rev. **105** (1957), 1127.
2) J. D. Jackson, Phys. Rev. **106** (1957), 330.
C. Hayashi, et al., Prog. Theor. Phys. **17** (1957), 651.
T. H. R. Skyrme, Phil. Mag. **2** (1957), 910.
Ya. B. Zel'dowich and A. D. Sakharov, J. E. T. P. **32** (1957), 947.
3) E. Teller, Zeit. für Physik **61** (1930), 458.
4) S. Cohen, et al., Phys. Rev. **110** (1958), 1471: Their method of calculation seems very similar to ours, though its details are not yet published.

Absorption Effects in Antinucleon Phenomena, II

Y. Kakudo, T. Kammuri and R. Nakasima*

Department of Physics, Osaka University, Osaka
**Research Institute for Atomic Energy, Osaka City University, Osaka*

September 15, 1958

In the previous note,[1] the absorption effects in antinucleon production processes within a nucleus were considered and it was found that the experimental results can be reproduced fairly well in the case

of antiproton production by taking account of the effects of the intranuclear motion and of the reabsorption of produced antiprotons. In the case of antineutron production by charge exchange of antiproton, however, desired agreement with the experiments was not obtained by taking into account only the absorption and reabsorption effects. As an attempt to remove the discrepancy, we use an improved nuclear model in the present note. In particular, since one may expect that the strong annihilation of antinucleon makes the reaction occur near the nuclear surface only, it is hoped to take more realistic shape of the nuclear surface.

According to many experimental results,[2] the form of charge distribution in a nucleus can be represented by Fermi-type-distribution such as

$$\rho(r)=\rho(0)/[\exp\{(r-c)/a\}+1], \quad (1)$$

$$c=1.08\,A^{1/3}\times 10^{-13}\ \text{cm},$$

$$a=0.53\times 10^{-13}\ \text{cm},$$

here the values of c and a are those determined by the electron scattering experiment on the charge distribution, but we assume that the expression (1) is applicable also to the density distribution in a nucleus. Therefore, $\rho(0)$ should be determined not as $\int\rho(r)dV=Ze$ but $\int\rho(r)dV=A$. Taking the diffused surface nuclear model, we express cross section of antineutron production as

$$\sigma=2\pi\sigma_{ex}\int_0^\infty bdb\int_{-\infty}^{\infty}\rho_p(r)dx \times\exp\left[-\sigma_a\int_{-\infty}^{\infty}\rho(r)dx\right], \quad (2)$$

where σ_{ex} is the elementary cross section for charge exchange process, σ_a is the annihilation cross section of antinucleon, $\rho_p(r)$ is the density of proton and x is the path length at the distance of b from the center of the nucleus. The expression (1) is well approximated by[3]

$$\rho(r)=\rho(0)u(r)$$

$$u(r)=1 \quad \text{for } 0\leq r\leq c-t/2,$$

$$u(r)=\frac{1}{2}-\frac{3}{2}\left(\frac{r-c}{t}\right)+2\left(\frac{r-c}{t}\right)^3 \quad \text{for } c-t/2\leq r\leq c+t/2, \quad (3)$$

$$u(r)=0 \quad \text{for } c+t/2\leq r,$$

$$t=4.23\times 10^{-13}\ \text{cm}.$$

Then $\rho(0)$, density at $r=0$, is given by

$$\rho(0)=A/[4\pi(c^3/3+ct^2/20)]. \quad (4)$$

In the integration on b of eq. (2), it is found that almost all the contribution comes from the region $c-t/2\leq b\leq c+t/2$, thus showing the importance of the diffuseness of nuclear surface. The results for carbon and lead targets, together with our previous one[1] based on model-A in which protons and neutrons in a nucleus are distributed uniformly within the same region, are tabulated as follows.

Table 1

	Present results	Previous results (Model-A)	Experimental results
σ for carbon	$0.74\times\sigma_{ex}$	$0.22\times\sigma_{ex}$	$1.33\left\{\begin{matrix}+0.54\\-0.50\end{matrix}\right.\times\sigma_{ex}$
σ for lead	$0.94\times\sigma_{ex}$	$0.20\times\sigma_{ex}$	$1.30\left\{\begin{matrix}+1.37\\-0.87\end{matrix}\right.\times\sigma_{ex}$

As is seen from this table, the results of the present calculations seem to be in better agreement with the experiments compared with the model of rectangular density distribution used previously. Furthermore, the effect of finite range of anti-nucleon-nucleon interaction, which was attempted by Agnew et al.[4] in interpreting the antiproton reaction cross section, may lead to an improved result. Therefore, one may expect that the process of anti-neutron production within a nucleus can be interpreted in terms of the absorption effects only.

After the completion of this calculation, we came to know that the similar calculations have been made by R. C. Weingart.[5] In his calculations, the values of $\bar{p}p$, $\bar{p}n$, $\bar{n}p$, and $\bar{n}n$ attenuation cross sections were used as the cross section responsible for absorption and it was assumed to $\sigma_a(\bar{p}p)=\sigma_a(\bar{p}n)=\sigma_a(\bar{n}p)=\sigma_a(\bar{n}n)=104$ mb, while in our calculation they are not the attenuation cross sections but the annihilation cross sections and the values are $\sigma_a(\bar{p}p)=\sigma_a(\bar{n}p)=89$ mb and $\sigma_a(\bar{p}n)=\sigma_a(\bar{n}n)=74$ mb. The different choice of these values is slightly reflected in final results. We would like to thank Dr. G. Goldhaber for sending the unpublished report of Weingart.

1) Y. Kakudo, T. Kammuri and R. Nakasima, Prog. Theor. Phys. **20** (1958), 243.
2) R. Hofstadter, Rev. Mod. Phys. **28** (1956), 214.
3) J. W. Cronin, R. Cool and A. Abshian, Phys. Rev. **107** (1957), 1121.
4) Agnew, Chamberlain, Keller, Mermod, Rogers, Steiner and Wiegand, Phys. Rev. **108** (1957), 1545.
5) R. C. Weingart, thesis, UCRL-8025 (1957), unpublished.

The Meissner-Ochsenfeld Effect in the Bogoliubov Theory

J. M. Blatt and T. Matsubara

The F. B. S. Falkiner Nuclear Research and Adolph Basser Computing Laboratories, School of Physics, The University of Sydney, Sydney, N. S. W.*

September 30, 1958

In the calculation of the magnetic behaviour of a system the quantity $S_{\mu\nu}(q)$ appears; in the limit of zero temperature

$$S_{\mu\nu}(q) = -\frac{2V}{c^2}\sum_k \frac{\langle 0|j_\mu(q)|k\rangle\langle k|j_\nu(-q)|0\rangle}{E_0-E_k} \quad (1)$$

where $j_\mu(q)$ is the q'th Fourier component of the current density operator. In the Bogoliubov theory[1] the operator $j_\mu(q)$ has matrix elements to double Fermi-type excitations. However, the Bogoliubov theory also contains Bose-type excitations with an energy spectrum

$$\varepsilon_c(p)=\frac{v_F|\vec{p}|}{\sqrt{3}} \quad (2)$$

where v_F is the velocity of an electron at the Fermi surface. The latter appear as "collective excitations" in the Bogoliubov approach, and are therefore missed by a perturbation expansion of the type used by Wentzel.[2] Without inclusion of Bose-type excitations, the sum (1) yields a result, $S_{\mu\nu}^{(F)}(q)$, with the property[3]

$$\sum_{\mu,\nu}\frac{q_\mu q_\nu}{q^2}S_{\mu\nu}^{(F)}\to 0. \quad (q \text{ small}) \quad (3)$$

* Also supported by the Nuclear Research Foundation within the University of Sydney.

This therefore violates the Buckingham identity[4]

$$\sum_{\mu,\nu} \frac{q_\mu q_\nu}{q^2} S_{\mu\nu}(q) = L = \frac{ne^2}{mc^2} \qquad (4)$$

Bogoliubov's expressions for the collective excitations show that there is a non-zero matrix element of $j_\mu(q)$ to the Bose-type state with $p=q$. Inserting this one extra term into (1), we get the additional contribution :

$$S^{(B)}_{\mu\nu}(q) = \frac{q_\mu q_\nu}{q^2} L \quad (q \text{ small}) \qquad (5)$$

Adding this term to (3), we find that (4) is satisfied, i. e., the result is gauge invariant. In the limit of small q the Fermi contribution to the kernel $K(q)$ is regular;[3] the leading term is the Bose contribution from (5), namely

$$K^{(B)}(q) = -L/q^2. \quad (q \text{ small}) \qquad (6)$$

This corresponds precisely to the London phenomenological equation, with a number n_s of "superconducting electrons" equal to the total number of electrons.

The spectrum (2) breaks off when $\mathcal{E}_c(p)$ exceeds 2Δ, where Δ is the energy gap in the Fermi-type excitation spectrum. The maximum value of p determined from this condition defines a wave number μ of the order of $(10^{-4}\,\mathrm{cm})^{-1}$. For $q > \mu$, the contribution (6) ceases, and $K(q)$ reduces essentially to the Fermi contribution. The length μ^{-1} is of the right order of magnitude for the coherence length of Pippard.[5]

It is interesting to compare the Bogoliubov theory with the quasi-chemical equilibrium, or pair correlation, approach to superconductivity.[6] In the latter, there is an energy spectrum $\mathcal{E}_k$ for single particles, and a spectrum $\gamma_j(p)$ for correlated pairs with centre-of-gravity momentum p. Both spectra are temperature dependent. The $\gamma_j(p)$ is related to the spectrum u_σ of the "quenched correlation matrix" by

$$u_\sigma = \exp[\beta(2\mu - \gamma_\sigma)]. \qquad (7)$$

We have found the following analogy useful : the Fermi-type excitations are analogous to single particles, the Bose-type excitations to correlated pairs with non-zero centre-of-gravity momentum ; hole theory is used for the single particles. The analogy is not perfect : the Fermi type excitations are superpositions of holes and electrons, not simply electrons ; the Bose-type excitations involve operators which create pairs of holes, and which scatter holes and electrons, as well as operators which create electron pairs. Nevertheless, our result can be understood physically in terms of this analogy. The dependence of (5) on q_μ is identical to that of the ideal Bose gas in the limit of zero temperature,[7] and arises in the same way. However, the coefficient involves not the number of correlated pairs in the ground state, but the total number of electrons. This is related to the energy gap in the single-particle spectrum, which prevents the single particles from making any significant contribution to the sum (4) in the limit of small q.

It should be noted that excitations without an energy gap are the essential factor in obtaining a gauge-invariant Meissner effect. If, as Anderson claims,[8] the Coulomb interaction between electrons destroys the spectrum (2) and reduces all the collective modes to high-lying plasma quanta, it would be difficult to see how a Meissner effect could result from the theory.

The specific heat contributed by (2) is a factor 10^9 below the ordinary lattice

specific heat. Thus the Bogoliubov theory leads naturally to a two-fluid model of superconductivity: the Fermi-type excitations give the "normal fluid", which appears in the specific heat but does not superconduct; the Bose-type excitations provide the "superfluid", which contributes essentially nothing to the specific heat, but is primarily responsible for the spectacular electromagnetic properties of superconductors.

We are grateful to Drs. S. T. Butler and M. R. Schafroth and Mr. R. M. May for valuable discussions concerning this work. One of us (T. M.) would also like to thank Prof. H. Messel and the Nuclear Research Foundation within the University of Sydney for making possible his stay in Sydney.

1) N. Bogoliubov, V. Tolmachev, and D. Shirkov "*A New Method in the Theory of Superconductivity*" Moscow (1958).
2) G. Wentzel, "*The Meissner Effect*", to be published.
3) R. M. May and M. R. Schafroth, private communication.
4) M. J. Buckingham, Nuovo Cimento, **5** (1957), 1763.
5) A. B. Pippard, Proc. Roy. Soc., **A191** (1947), 370, 385, 400, **A203** (1950), 98, 195, 210.
6) M. R. Schafroth, S. T. Butler, and J. M. Blatt, Helv. Phys. Acta, **30** (1957), 93.
7) V. L. Ginzburg, Usp. Fiz. Nauk., **48** (1952), 25; M. R. Schafroth, Phys. Rev. **100** (1955), 463.
8) P. W. Anderson, "*The Random Phase Approximation in the Theory of Superconductivity*", to be published.

Errata

A Comment on Bardeen's Theory of Superconductivity

S. Nakajima and T. Kasuya
Prog. Theor. Phys. **18** (1957), 662
(October 3, 1958)

The letter published by us under the above title contains incomplete and ambiguous points. The matrix elements of H_2 oscillates in sign and no gain in energy results from the trial function proposed in that letter. However, there still remains a possibility of obtaining condensed state if we consider that the magnitude V of the Hamiltonian H_2 is constant. Actually, however, V is proportional nearly to $(|C|^2\hbar\omega_0 - 4\pi e^2/k_c^2)$ where $|C|$, ω_0 and k_c mean respectively the coupling constant of electron phonon interaction, Debey cut off angular fequency of phonon and the shielding constant of Coulomb interaction, and k_c becomes very large in condensed state. Therefore V becomes negative and the condensed state destroyed again. From these considerations our comment concerning to Bardeen's theory is not correct.

JOURNAL OF THE PHYSICAL SOCIETY OF JAPAN
Volume 13, Number 9, September 1958

CONTENTS

SHORT NOTES

Errata